

基于生成对抗网络的雾霾场景图像转换算法

肖进胜 申梦瑶 雷俊锋 熊闻心 焦陈坤

(武汉大学电子信息学院 武汉 430072)

摘要 本文提出了一种新的基于生成对抗网络的雾霾场景图像转换算法,生成对抗网络 GAN 作为无监督学习的方法,无法实现图像像素与像素之间映射,即生成图像不可控.因此,基于模型的加雾算法存在参数不确定性和应用场景局限性,本文提出了一种新方法的新应用,利用生成对抗网络实现图像转换.该方法基于生成对抗网络 GAN 模型,改进了 GAN 的生成器和判别器,进行有监督学习,以训练雾霾图像生成像素与像素之间的映射关系,实现无雾图像与有雾图像之间的转换.以图像加雾为例,本文分别设计了生成网络和判决网络,生成网络用于合成有雾图像,判决网络用于辨别合成的雾霾图像的真伪.考虑到雾霾场景图像转换的对应效果,设计了一种快捷链接沙漏形生成器网络结构,采用无雾图像作为生成网络输入,并输出合成后的有雾霾图像;具体来看,将生成网络分成编码和解码两部分,并通过相对应间隔的卷积层来保留图像的底层纹理信息.为了更好地检验合成雾霾图像的真实程度,设计了漏斗形全域卷积判决器网络,将合成图像和目标图像分别通过判决器辨别真伪,采用全域卷积,利用神经网络进行多层下采样,最终实现分类判决,辨别图像风格.此外,本文提出了一种新的网络损失函数,通过计算 GAN 损失和绝对值损失之和,以训练得到更为优秀的图像转换结果.GAN 损失函数的作用是使生成对抗网络 GAN 模型训练更加准确,而雾霾图像合成算法实际上是一个回归问题而非分类问题,生成器的作用不仅是训练判决器更加灵敏,更重要的是要生成与目标图像相似的图像.因此利用优化回归问题的绝对值损失函数,作用是为了准确学习像素间的映射关系,避免出现偏差和失真.最后本文对多类不同图像进行图像的雾霾场景转换并进行评估,分别测试该算法的图像加雾和去雾效果,并与其他算法进行对比测试.对于加雾效果,在合成场景、虚拟场景下,与软件合成效果进行对比,本文算法效果明显比软件合成效果好,不会出现色彩失真;在真实场景下,本文算法与真实拍摄的雾霾天气进行对比,结果十分相近;并且与其他 GAN 图像转换算法进行对比,本文算法具有明显的优势.同样本文算法在去雾效果上优势也十分明显.结果表明,本文所提基于生成对抗网络的雾霾场景图像转换算法,在主观效果和客观指标上均具有明显优势.

关键词 图像处理;图像转换;雾霾场景;生成对抗网络;深度学习

中图法分类号 TP311

DOI号 10.11897/SP.J.1016.2020.00165

Image Conversion Algorithm for Haze Scene Based on Generative Adversarial Networks

XIAO Jin-Sheng SHEN Meng-Yao LEI Jun-Feng XIONG Wen-Xin JIAO Chen-Kun

(School of Electronic Information, Wuhan University, Wuhan 430072)

Abstract In this paper, synthesis of hazy image based on generative adversarial net is described in detail. As an unsupervised learning method, the generative adversarial networks cannot learn the mapping between image pixels, which means the generation of images is uncontrollable. Therefore, because of the uncertainty of the parameters of the algorithms and the limitation of the application scenarios, a new application of the new method is proposed in this paper. The networks can learn the mapping between input and output image, and learn a loss function to

收稿日期:2019-01-28;在线出版日期:2019-09-11. 本课题得到国家重点研发计划(2016YFB0502602)、国家自然科学基金(61471272)和湖北省高等学校省级教学研究项目(2017016)资助. 肖进胜,博士,副教授,中国计算机学会(CCF)会员,主要研究方向为计算机视觉、视频图像处理. E-mail: xiaojis@whu.edu.cn. 申梦瑶,硕士研究生,主要研究方向为计算机视觉. 雷俊锋(通信作者),博士,副教授,主要研究方向为视频图像处理、计算机视觉. E-mail: jflei@whu.edu.cn. 熊闻心,硕士研究生,主要研究方向为数字图像处理. 焦陈坤,硕士,主要研究方向为计算机视觉.

train this mapping. The algorithm is based on GAN, and the improved generator and discriminator are proposed. And supervised learning is carried out to train the mapping between pixels in hazy images and haze-free images. Taking synthesis of hazy image as an example, we propose two networks: the one is generative network, which is used to generate hazy images, and the other one is discriminative network, which is used to identify the images. Here, considering the corresponding effect of image conversion, a hourglass-shaped fast link generator network is designed, which uses the fog-free image as the input. Specifically, the generated network is divided into two parts, encoding and decoding, and the underlying texture information of the image is preserved by adding corresponding convolution layers. Then, a funnel-shaped global convolutional judge network is designed to test the results of smog image synthesis. The composite image and the target image are respectively identified by the decider. The global convolutional neural network is used for multi-level downsampling to achieve classification and discerning image style. Meanwhile, the loss function of network is also modified. By calculating the sum of GAN loss and ABS loss, better results can be gotten. The loss function of GAN is to make the training of GAN model more accurate. However, the hazy image synthesis algorithm is actually a regression problem rather than a classification problem. And the generator is tasked not only to fool the discriminator but also to be near the ground truth Output. So in order to learn the mapping accurately and avoid deviation and distortion, the using of the ABS loss function is helpful. Finally, we analyze and contrast the experimental results carefully. This paper evaluates the transformation of different haze scenes, tests the effect of our hazing and dehazing algorithm, and compares it with other algorithms. For hazing effect, in synthetic scene and virtual scene, compared with the effect of software, the effect of our algorithm is obviously better than that of others, and there is no color distortion. In real scenes, the results of our algorithm are very similar to those of the real foggy and hazy scene. What's more, compared with other GAN image conversion algorithms, our algorithm has obvious advantages. Similarly, it can be seen that the effect of our haze removal task is also very obvious. Experiments demonstrate that the proposed algorithm has better performance than state-of-the-art methods on both synthetic and real-world images qualitatively and quantitatively.

Keywords image processing; image conversion; haze scene; generative adversarial net; deep learning

1 引言

雾不仅仅是一种天气现象,它是美学摄影中的重要因素,也是虚拟现实中的重要场景,还是道路交通的梦魇。在智能驾驶、城市监控、卫星遥感领域雾会使图像失真。去雾成为图像处理部分的重要内容之一。目前利用深度学习进行去雾是主流研究方向。但该方法需要有雾和无雾匹配图片对作为训练集。由于时间、地点、相机参数的限制,以及图像的匹配及变换算法使得大量图像对的采集异常困难。因此考虑通过算法人工地在图像上合成雾霾。传统雾霾合成方法主要是基于模型的算法。近些年来,深度学习作为当今人工智能领域最热门的方向,在图像识别、

语义分割、自然语言处理等领域取得了重大突破,显示出强大的学习和处理能力。因此,一些研究人员将深度学习应用于图像生成^[1]任务中,并在图像风格化方面取得了很好的效果。综上,在雾霾场景中,图像转换方法可以分为两大类:基于模型和基于学习的方法。

1.1 基于模型的方法

由于图像中物体景深不同^[2],雾霾在不同深度有不同的效果,基于创建深度的模型如 Liu 等人^[3]提出了使用双目立体视觉进行照片编辑的雾模拟方法,首先利用 iSGM 计算出图像的景深信息,然后利用中值滤波器进行平滑,最后根据深度信息加雾。主要依据处理后的景深图与三种效应:光直接射入相机;光散射进入相机;其他光源的光散射进入相机。

还有一类基于模型的方法是基于大气光散射模型进行雾的合成。大气散射模型^[4]为: $I(x) = J(x) \cdot t(x) + A(1 - t(x))$, 其中 I 是观察图像, A 是空气光颜色矢量, J 是场景交互点处的表面辐射矢量中与像素 x 对应的真实世界射线。 $t(x)$ 被称为透射映射, 它表示在观察者和场景中的一个表面点之间的整个路径中存活而不被散射的光的相对部分。 Zhao 等人^[5]获得照明的近似值以及底层场景的几何信息, 通过生成输入图片的透射图和立体图, 结合输入照片、透射率图和体积图合成雾状图片。然而大气散射模型不能完全正确地描述雾天的退化机制且大气光参数难以选取, 该模型只能模拟单散射雾效应, 不能捕获体积焦散效应, 也不能很好地处理图像中的折射物体。

1.2 基于学习的方法

图像生成算法实际上是一幅图像到另一幅图像的转换。利用深度学习合成不同场景的图像来进行风格转换, 最初是由 Hertzmann 等人^[6]提出的, 其是采用非参数纹理模型训练图像对的泛用框架。 Gatys 等人^[7]提出了一种基于高性能卷积神经网络的风格转换算法, 用 CNN 学习到的图像深层特征分别表示图像的风格和内容; 其输出图片分辨率低、存在噪音。生成对抗网络 GAN^[8]采用对抗性损失理念, 使生成图像接近真实图像, 但无监督学习无法实现图像像素间的映射, 即生成图像不可控。 Isola 等人^[9]提出 pix2pix 算法使用条件 GAN (conditional GAN) 学习输入到输出的映射, 更适用于高度结构化的图形, 对其他图像会产生模糊化问题。 Tian 等人^[10]提出了一种新的图像合成框架, 将图像映射到特征及将潜在因素转换为图像, 适用于合成图像纹理结构, 对雾霾合成缺少泛化能力。基于模型的加雾算法往往存在参数不确定性和应用场景局限性, 由此, 本文提出了一种新方法的新应用, 利用生成对抗网络合成雾霾图像。改进了 GAN 生成器和判别器, 进行有监督学习, 以生成像素间的映射关系, 实现无雾图像与有雾图像之间的转换。

本文提出了基于生成对抗网络的雾霾场景图像转换算法, 提出了几点创新:

(1) 借鉴 U-net 网络结构提出改进的生成器网络, 设计了一种含跳层的沙漏形卷积神经网络。将生成网络分成编码和解码两部分, 通过相加对应间隔的卷积层来保留图像的底层纹理信息;

(2) 在判别器网络, 本文设计了一种漏斗形网络结构, 并采用全域卷积, 利用神经网络进行多层下采样, 最终实现分类判决, 辨别图像风格;

(3) 修改损失函数, 分别利用 GAN 损失函数和绝对值损失函数; GAN 损失函数作用是使生成对抗网络 GAN 模型训练更加准确。而雾霾图像合成算法实际上是一个回归问题而非分类问题, 因此利用优化回归问题的绝对值损失函数, 作用是为了准确学习像素间的映射关系, 避免出现偏差和失真。

本文第 2 节介绍生成对抗网络的相关工作; 第 3 节主要介绍本文基于生成对抗网络的图像生成算法, 描述算法总体框架的流程步骤, 给出算法中生成器和判决器的网络结构图以及损失函数; 第 4 节首先说明如何生成训练集和测试集, 然后对比分析本文算法中不同模块对模型性能的影响。最后针对本文算法效果, 做了大量的对比实验并且进行系统全面的分析与评价。评价主要从两方面入手: 主观视觉效果和客观评价指标; 第 5 节总结全文。

2 相关工作

下面将对生成对抗网络 GAN 的演变、如何利用 GAN 实现图像风格转换进行介绍。

2.1 生成对抗网络 GAN

生成对抗网络是一种无监督机器学习方法。生成对抗网络中有两个网络模型: 生成器和判别器。判别器判定样例是来自数据集还是生成器合成的图像, 生成器尽可能使生成图像以假乱真迷惑判别器使其难辨真伪, 两个网络模型相互对抗以提升各自算法能力直到判别器无法分辨合成图像与数据集图像。

生成器想要从真实图像 y 中学习其分布 p_g , 定义输入噪声变量 $p_z(z)$, 代价函数为 $V(D, G)$:

$$\min_G \max_D V(D, G) = E_{y \sim p_{\text{data}}(y)} [\log D(y)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (1)$$

z 表示输入生成器 G 的噪声, $G(z)$ 表示 G 生成的图片。 $D(y)$ 表示判别器 D 判断是真实图片的概率, E 为数学期望。 y 是真实的, 希望 $D(y)$ 越接近 1。 $D(G(z))$ 是 D 判断 G 生成图片真实的概率。 G 希望 $D(G(z))$ 尽可能大, 这时 $V(D, G)$ 会变小。 D 希望 $D(y)$ 越大, $D(G(z))$ 越小。作为无监督学习, 该方法无法实现像素间的转换。

2.2 DCGAN

DCGAN^[11]将卷积神经网络应用在 GAN 中, 以便对 GAN 学习到的滤波器进行定性分析并使训练过程趋于稳定。首先在生成器和判别器网络中用全卷积网络结构替换池化层和全连接层; 其次用批量标准归一化层解决糟糕初始化引起的训练问题,

使得梯度传播更深. 具体来看, 生成模型 G 以 100 维噪声向量 $\mathbf{Z}$ 作为输入, 每个反卷积层通道数减半, 图像尺寸加倍. 生成模型输出 $64 \times 64 \times 3$ 的 RGB 图像, 判决模型 D 输出是一个标量, 判别生成图像与真实图像. 该方法只能随机生成图片, 转换的结果不可控, 因此无法实现图像与图像的转换.

2.3 Pix2pix

同样将卷积神经网络应用在 GAN 上的还有 pix2pix. 该算法用于图像的像素间的转换, 提出了利用条件 GAN 在输入图像上生成相应的输出图像. 该方法的生成器采用 U-net 网络^[12], 判决器采用 patchGAN 网络, 即仅提取图像中某一块进行卷积操作. 该框架适用于多种应用场景, 可以很好地迁移, 应用多种转换操作, 例如黑白图像与彩色图像的转换、卫星图像与绘制地图的转换、手绘物体图像与真实物体图像的转换. 但是从图像细节较少向图像细节较多的方向转换时, 则会出现图像失真.

2.4 CycleGAN

CycleGAN^[13] 可利用非成对图像进行训练, 主要依赖于循环一致性损失, 该损失要同时学习正向映射即 $G: X \rightarrow Y$, 及反向映射 $F: Y \rightarrow X$, 并要求 $F(G(x)) \approx x$, $G(F(y)) \approx y$. 即将 X 转换到 Y 空间过程可逆. 循环一致性损失就定义为

$$L_{\text{cyc}}(F, G, X, Y) = E_{x \sim p_{\text{data}}(x)} [\|G(F(x)) - x\|_1] + E_{y \sim p_{\text{data}}(y)} [\|F(G(y)) - y\|_1] \quad (2)$$

再在生成对抗网络 GAN 中引用 GAN 损失. 正向映射也即 $G: X \rightarrow Y$, 定义其判决器为 D_Y , 则其 GAN 损失 $L_{\text{GAN}}(G, D_Y, X, Y)$ 为

$$L_{\text{GAN}}(G, D_Y, X, Y) = E_{y \sim p_{\text{data}}(y)} [\log D_Y(y)] + E_{x \sim p_{\text{data}}(x)} [\log(1 - D_Y(G(x)))] \quad (3)$$

定义反向映射 F 的判决器为 D_X , 由此可以同样定义 $L_{\text{GAN}}(G, D_X, X, Y)$. 最终的损失含三部分:

$$L = L_{\text{GAN}}(F, D_Y, X, Y) + L_{\text{GAN}}(G, D_X, X, Y) + \lambda L_{\text{cyc}}(F, G, X, Y) \quad (4)$$

该方法基于 pix2pix 修改损失函数; 图像训练集不需要成对数据所以应用范围更广泛. 但由于训练集不匹配, 无法在绝对映射关系中进行训练, 只能猜测真实的映射关系, 因此其学习到的映射关系会出现偏差, 导致训练结果出现非期望转换风格.

3 本文算法结构

通过对 GAN 发展的研究, 本文提出基于 GAN

的雾霾场景图像转换算法. 生成器借鉴 U-net 网络结构, 设计了一种含跳层连接的沙漏形卷积神经网络, 包含编码和解码两部分, 并通过相加对应间隔卷积层来保留图像的底层纹理信息. 判决器采用漏斗形网络结构, 并采用全域卷积进行多层下采样, 最终实现分类判决, 辨别图像风格; 修改损失函数, 以准确学习像素间的映射关系, 避免出现偏差和失真方法. 以雾霾场景生成为例, 算法流程图如图 1.

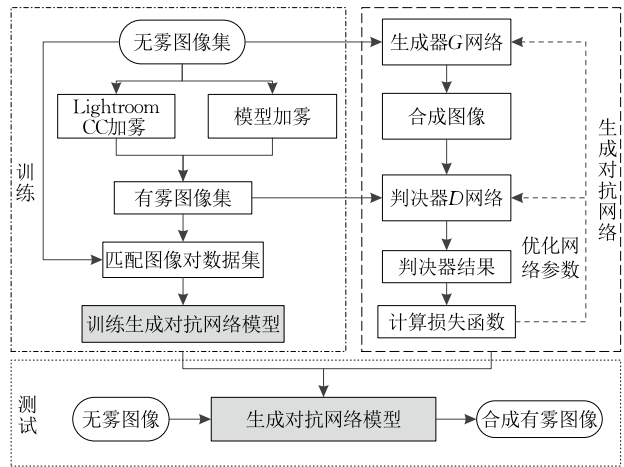


图 1 本文算法整体流程框图

本算法主要分训练和测试两个阶段. 在训练阶段将 GAN 网络模型进行优化, 在测试阶段输入无雾图像通过 GAN 网络模型得到输出图像. 之后通过生成网络和判决网络不断迭代, 优化网络参数.

3.1 生成器网络结构

利用多层卷积构建生成器网络在 DCGAN 中已有提出, 但仅用卷积层提取特征没有考虑生成图像与输入图像在纹理特征和图像结构上的联系, 造成细节模糊. 考虑以上缺陷, 本文改进生成器网络, 设计了一种含跳层的沙漏形卷积神经网络, 将原网络中的池化层替换成步幅为 2 的卷积操作进行下采样. 生成器网络分为编码器和解码器, 相对于 U-net 网络编解码层数更深, 网络输入尺寸逐层减小, 直到瓶颈层, 然后该过程再被反转.

然而在雾霾图像生成上, 实际输出图像与输入图像虽然视觉上不同, 但是都有相同的底层结构和边缘. 因此, 我们要考虑将输入图像中的结构进行一定的保留, 在输入和输出之间共享一部分信息, 并且希望将这些信息直接穿过网络不经过瓶颈层. 在卷积层中加入跳层, 形成如图 2 的网络结构.

该网络输入输出图像尺寸均为 $256 \times 256 \times 3$. 左侧编码器包含 8 层 4×4 卷积-Batch Normalization-

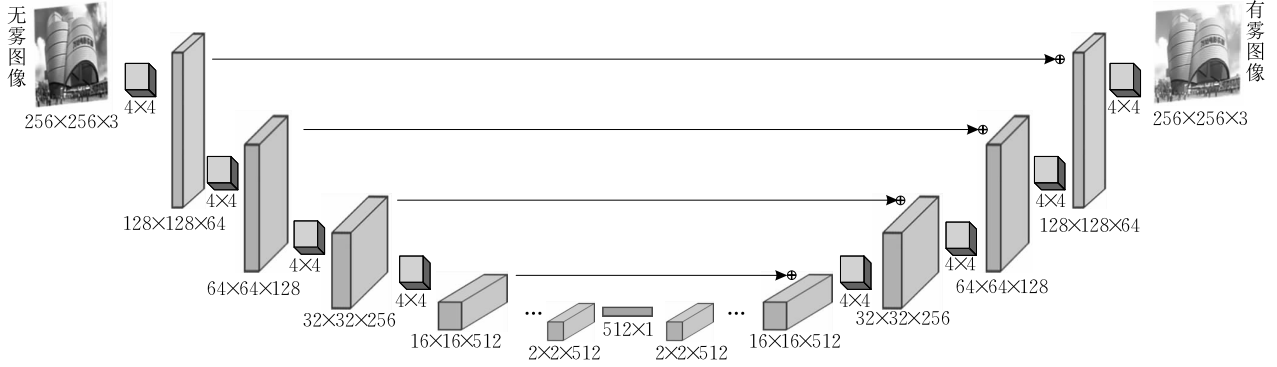


图2 生成器网络结构

prelu组合,每层利用 4×4 、步幅为2的卷积核进行下采样,输出一个一维向量。八层卷积个数为

$$64-128-256-512-512-512-512-512,$$

右侧解码器与编码器中心对称,含8层 4×4 反卷积-Batch Normalization-relu组合层,用来做上采样:

$$512-512-512-512-512-256-128-64,$$

解码器每层的结果来源于自身卷积结果与编码器对应卷积层相加的结果。故每层实际的卷积个数:

$$512-1024-1024-1024-1024-512-256-128.$$

卷积层后采用Batch Normalization^[14]使训练后每一层数据归一化且分布在一定范围内,减小数据大幅度波动以加快训练速度,提高网络泛化能力。

3.2 判别器网络结构

3.2.1 判别器网络设计

本文设计了一种漏斗形判别器网络结构,采用全域卷积实现多层下采样,最终辨别图像风格。区别于pix2pix对某一图像块进行卷积的patchGAN。本文采用全域卷积减小操作图像尺寸,图像尺寸越大可以进行更深地下采样,提取更多的图像特征,从而保证细节判决更为准确。网络结构如图3所示。

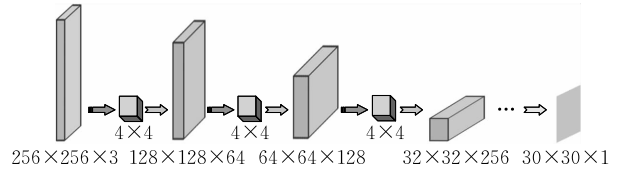


图3 判别器网络结构

判别器前 n 层是下采样层,最后判决层输出一维向量。下采样层采用 4×4 、步幅为2卷积核及卷积-Batch Normalization-prelu组合。每经过一个下采样层,图像的长宽分别为原来的一半。第一个下采样层的卷积核个数设为64,后每层卷积核数设为上一层的2倍。由于最终判决层输出一维向量,因此下采样层最大为6。下采样层数 n 一定程度上影响图像生成结果,第3.2.2节将对 n 进行进一步分析。

3.2.2 判别器下采样层 n 的影响

理论上下采样卷积层越深提取的特征越精确,但限于输入图像尺寸的限制 256×256 ,每经过一次下采样图像长宽尺寸均缩小一半,由于最终判决层输出一维向量,因此下采样层最大为6。但下采样层越多,参数越多,计算复杂度越大,耗时越多。因此,本文分别对下采样层 $n=3,4,5,6$ 在迭代20000次时得到的结果进行主观对比,如图4所示。



图4 “草原”图像对比结果

由图4可见,虽然草原天空在不同下采样层数下效果不同,但整体上加雾效果影响很小,各输出图像都能很好的保持图像的细节和边缘。考虑到计算量,本文实验结果的下采样层个数 n 设定为3。

3.3 loss函数

3.3.1 loss函数定义

本算法基于GAN解决生成图像即回归问题。GAN损失函数 $loss_{GAN}$ 使判别器和生成器更加准

确;回归损失函数 $loss_{ABS}$ 使生成图像与目标图像更加接近. 故代价函数 $loss$ 由这两部分组成:

$$loss=loss_{GAN}+\lambda loss_{ABS} \tag{5}$$

其中 λ 为可调参数. 本文的生成器 G 网络目的是学习输入图像 x 到目标图像 y 的映射, $G: x \rightarrow y$, 使生成图像 $G(x)$ 与目标图像 y 十分接近, 难辨真假; 判决器 D 网络的目的是尽可能判断出图像是生成图像. 因此 $loss_{GAN}$ 定义如下:

$$loss_{GAN}(G,D)=E_{y \sim p_{data}(y)}[\log D(y)] + E_{x \sim p_{data}(x)}[\log(1-D(G(x)))] \tag{6}$$

本文算法虽然基于 GAN 生成对抗网络, 但本质上还是一个回归问题, 因为最终目的是生成图像,

因此引入回归问题的损失函数, $loss_{ABS}$ 定义如下:

$$loss_{ABS}=E_{x,y \sim p_{data}(x,y)}[\|y-G(x)\|] \tag{7}$$

最终目标是使得训练后的生成器 G^* 满足下式:

$$G^*=\arg \min_G \max_D loss_{GAN}+\lambda loss_{ABS} \tag{8}$$

3.3.2 $loss$ 函数中的参数 λ

该参数的调整主要是对两个 $loss$ 函数权重的考量. 虽然本算法基于 GAN 生成对抗网络原理, 但是算法目的是生成图像, 也即回归问题. $loss_{GAN}$ 的目的是使判决器和生成器更加准确, 而 $loss_{ABS}$ 的目的是使生成图像与目标图像更加接近. 于是考虑将 λ 分别取 100 和 0, 当 $\lambda=0$ 时也即 $loss$ 函数只有 $loss_{GAN}$, 没有 $loss_{ABS}$ 的情况. 实验结果如图 5 所示.



图 5 不同 $loss$ 函数对结果的影响

由图 5(b) 和 (c) 的局部放大区域可以明显看出, $loss$ 函数定义为 $loss_{GAN}$ 时会出现块效应, 也即图像中墙体区域出现成块均匀的条纹, 使图像失真; 当 $loss$ 函数定义为 $loss=loss_{GAN}+\lambda loss_{ABS}$, 加雾结果的块效应明显减弱甚至消失, 提升了图像的纹理清晰度. 在参数 λ 大小的实验对比中, 从图 6 中损失函数的变化可以看出, GAN 损失和回归损失不在一个数量级上, 需要适当增大回归损失的影响而不影响 GAN 损失发挥作用. 因此通过增大 $loss_{ABS}$ 权重来使生成图像与目标图像更加接近. 分别设置 λ 为 1,

10, 30, 50, 70, 90, 100, 在 20 000 次训练迭代结果下, 进行主观对比. 结果如图 7 所示.

由图 7 可以看出, λ 较小时, 天空区域出现明

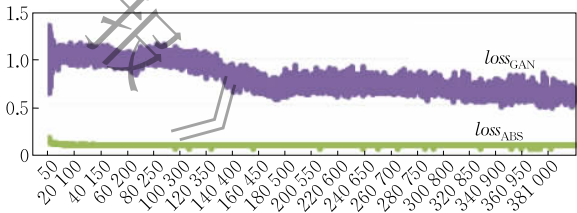


图 6 $loss$ 值的变化

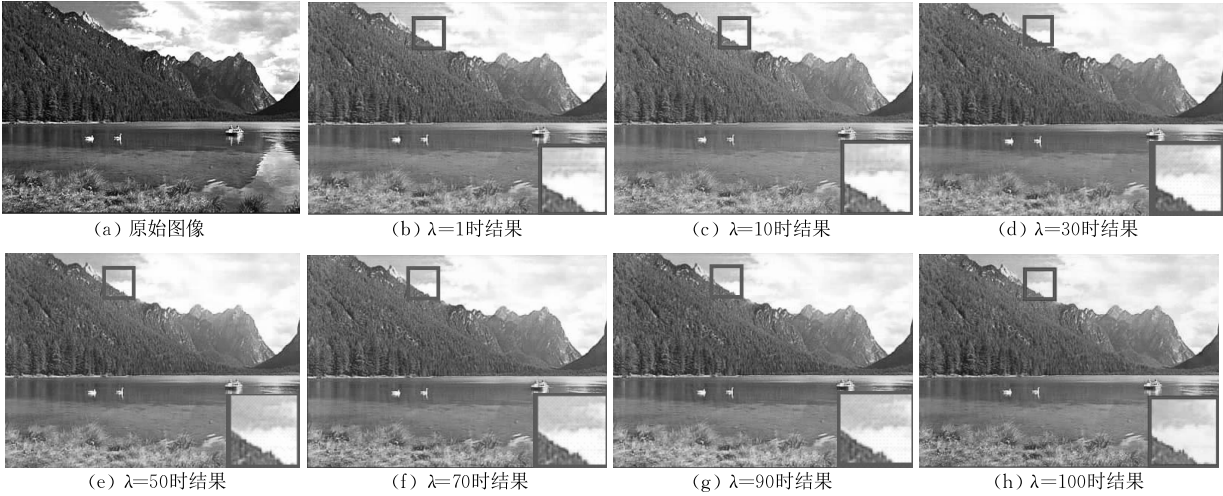


图 7 不同参数 λ 对实验结果的影响

显的块效应；随着 λ 增大，块效应明显减弱甚至消失。故在本文的实验结果的 $loss$ 函数中参数 λ 选为 100。

4 实验结果及分析

该算法可以同时实现图像加雾和反向图像去雾。限于篇幅，选取图像加雾方向和相关算法进行对比。本文算法实验环境详细信息如下：

硬件设备：CPU: Intel Core i7-5820K@3.30 GHz x 12; GPU: NVIDIA GeForce TITAN X; 内存: 16GB;

软件配置：操作系统为 64 位 ubuntu 14.04 LTS; MATLAB R2014a; CUDA Toolkit 7.0; OpenCV 3.0.

4.1 训练和测试集

本文算法使用 Tensorflow 框架，训练每迭代 100 次耗时 11.2 s。本文利用软件 Adobe lightroom CC 的加雾功能对 Middlebury Stereo Datasets 和网上收集的无雾图像进行加雾。采用软件加雾作训练集的原因在于其能快速高效地模仿雾霾场景，且大部分与真实有雾场景十分接近；但软件加雾在不同内容图像处理时会出现特定失真，使雾霾呈现绿色或蓝色，结果见 4.3.1 节。此问题可通过 GAN 网络优化迭代来规避。

Pix2pix 算法实验中采用 400 对匹配图像对作为训练集。鉴于此，本文利用软件加雾分别对 76 张

无雾图像加软件浓度参数为 30, 40, 50, 60, 70, 80, 90, 100 的雾，其中 50 张室内图像来自 Middlebury Stereo Datasets, 26 张室外无雾图像收集自网上，最终形成 608 对含不同浓度雾的有雾与无雾图像的匹配对作为训练集。真实图像测试集同样选择 Middlebury Stereo Datasets 和网上收集的无雾图像共计 45 张，拍摄图像测试集本文共采集并成功配准 20 对。

4.2 迭代次数的选择

每 50 次迭代记录 $loss$ 、 $loss_{GAN}$ 和 $loss_{ABS}$ ，如图 7 所示。由变化趋势可以看出，在 160 000 次左右 $loss_{GAN}$ 达到局部最小值，之后缓慢减小是判别器继续在训练以便判断更加准确； $loss_{ABS}$ 在 0.07 附近波动是因为在 20 000 次之后生成图像与目标图像已基本接近，故其趋于稳定。本文目的是生成图像，而非优化判别器的分类能力，因此选择 20 000 次迭代训练的网络。

4.3 实验结果

4.3.1 图像加雾主观结果分析

(1) 真实图像加雾

我们对网上收集的无雾图像进行雾霾生成实验，且与加雾软件的加雾结果作对比。软件加雾图像由 Adobe lightroom CC 加雾参数为 50 来得到。并且我们将本文算法得到的加雾图像利用 He^[15] 去雾经典算法 DCP 得到去雾图像，对比结果如图 8、图 9。

由对比图像可以看出，本文算法得到的雾霾合成



图 8 “丛林”图像对比结果



图 9 “秋天”图像对比结果

效果与软件加雾效果十分相似且十分逼真,如图 8;图 9 中软件加雾效果略有色彩失真,图像整体偏蓝,更像“烟雾”效果,而本文算法合成的雾霾图像仍然保持图像整体的色彩风格.从去雾效果看,本文算法合成的雾霾图像经过经典去雾算法后,结果与原始图像十分接近.从图(d)与图(a)对比可见,本文算法合成的雾霾图像能够保留细节,使去雾算法能够很好地应用于合成雾霾图像,让去雾图像恢复出很多图像细节.在 4.3.2 节我们也会从客观指标上证明这一点.

(2) 虚拟图像加雾

与 Tarel^[16]中的虚拟合成图像及其加雾对比结果如图 10、图 11. Tarel 加雾图像是利用 SiVIC 软件生成虚拟图像,再基于 Koschmieder 法则以能见度 80m 加雾.由结果可以看出, Tarel 的加雾更加适合浓雾、团雾效果,而本文算法则是重点在雾和霾生成.比起团雾、浓雾效果,雾霾在空气中分布更加均匀,这个结论也能从上面两个对比结果看出.

(3) 拍摄图像加雾

对手机拍的真实雾霾天气图像进行对比实验.手机型号为魅族 MX5,图像由后置摄像头采集.

由图 12、图 13 可以看出,本文算法的雾霾合成结果与雾霾天拍摄结果十分接近,在天空、高楼及树叶边缘没有颜色失真,且没有过亮和过暗的区域出现.而图 13 中软件加雾效果和真实场景相差较多.

(4) GAN 相关算法对比

将本文算法与相关 GAN 算法进行对比.首先, DCGAN 对 CNN 网络进行修改以便与 GAN 结合,但无监督学习使生成内容不可控,而雾霾场景转换要求图像内容不变,因此 DCGAN 无法实现图像的加去雾.于是将本文算法与 pix2pix、DRPAN^[17]和 cycleGAN 在相同训练集下得到的结果进行对比.

由图 14、图 15, cycleGAN 在图像细节和色彩方面都不能够得到很好的加雾效果,如图 14(b)中的玩偶肚皮区域出现明显失真;对于 pix2pix 算法,



(a) 虚拟无雾图像 (b) Tarel加雾结果 (c) 本文算法

图 10 虚拟图像 1 对比结果



(a) 虚拟无雾图像 (b) Tarel加雾结果 (c) 本文算法

图 11 虚拟图像 2 对比结果



(a) 真实无雾图像 (b) 真实雾天图像 (c) 软件加雾图像 (d) 本文算法加雾图像

图 12 拍摄图像“广场”对比结果

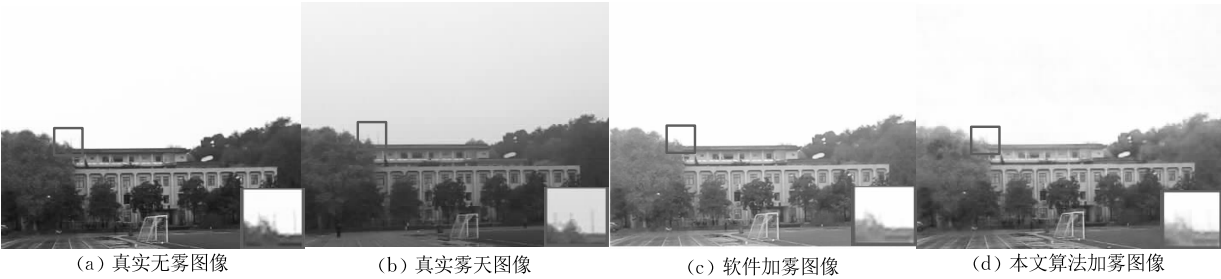


图 13 拍摄图像“操场”对比结果

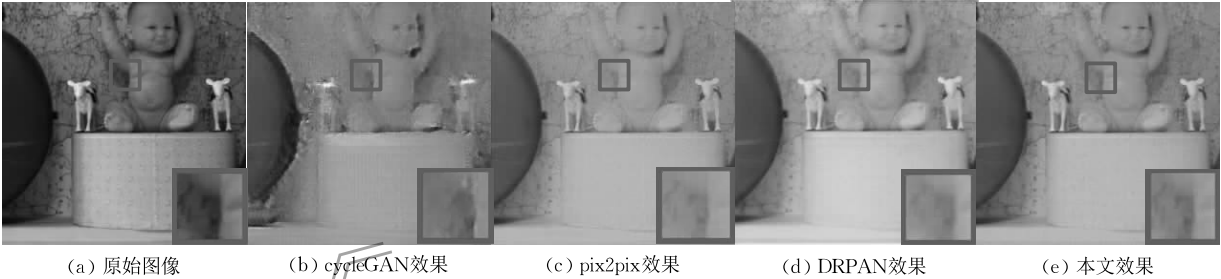


图 14 其他 GAN 算法对比结果 1

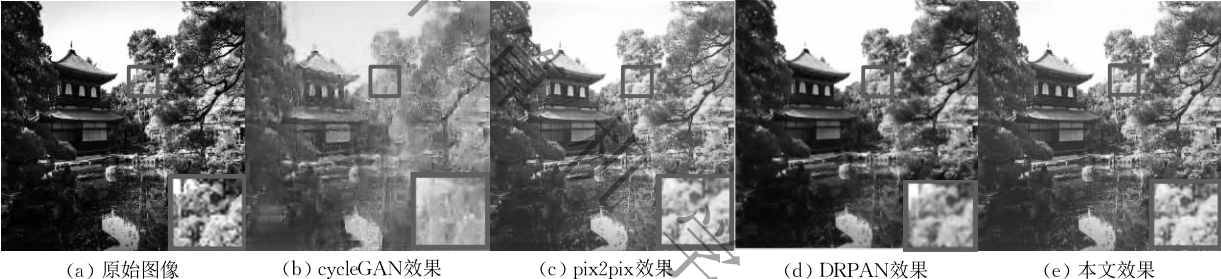


图 15 其他 GAN 算法对比结果 2

同样能够实现加雾效果,但对于细节处理不够好,与背景色彩十分相近,细节丢失;对于 DRPAN 算法,对于细节的处理不够细致,比如地图的背景、线条十分模糊,而本文算法则没有这种情况;再如图 15(b)中 cycleGAN 色彩失真,图像内容模糊,而 pix2pix 中框内区域的有些树叶会泛白,DRPAN 的效果并没有明显的雾霾效果,只是色彩比原始图像略有不同.而本文算法都能够很好地处理细节,加雾效果明显,并且没有色彩失真.

4.3.2 图像加雾客观指标对比

Choi 等人^[18]提出了一个计算雾浓度(fog density)的算法,利用该算法我们对原始图像和本文算法雾霾合成结果分别求出雾浓度指标,结果见表 1,包含

4.3.1 节中的真实图像、拍摄图像.由表 1 可以看出,本文算法输出结果比原始图像大,Fog density 数值越大证明图像的雾越浓.可见,本文算法能够在原始图像上有效地在一定范围内加雾.

其次,本文参考其他典型基于学习进行图像去雾文章,在进行客观指标评价时采用 PSNR 和 SSIM 来评判生成图像的效果,因此本文同样采用 PSNR 和 SSIM 客观指标对进行真实图像定性对比分析.对比结果如表 2 所示,包含 4.3.1 节中的真实图像.将原始图像作为基准图像,分别对本文算法加雾结果、将本文算法得到的加雾图像利用 DCP 算法得到去雾图像求 PSNR 和 SSIM 指标.

表 1 fog density 指标对比

Fog density	图 9	图 10	图 11	图 12	图 15	图 16
原始图像	0.1325	0.1798	0.1353	0.2056	1.3726	2.3196
本文雾霾合成结果	0.3959	0.6358	0.3607	0.5969	2.2680	3.1868

表 2 PSNR 和 SSIM 客观指标对比

	本文雾霾合成结果		DCP 去雾结果	
	PSNR	SSIM	PSNR	SSIM
图 9	13.6483	0.7526	24.4476	0.9253
图 10	13.3327	0.6690	22.1290	0.8938
图 11	14.7933	0.8268	22.1540	0.9427
图 12	13.8534	0.7715	23.7279	0.9153

由表 2 的客观指标可知,对于 PSNR 和 SSIM 指标,本文雾霾合成结果使得 PSNR 和 SSIM 明显下降,SSIM 在 0.7 左右,PSNR 在 14 左右,可以证明本文算法能够使图像对比度降低,这也是雾霾天气对图像质量影响的共识. 经过本文算法处理得到的雾霾图像再进行去雾,图像的 PSNR 和 SSIM 明显提升,SSIM 在 0.9 左右,PSNR 在 20 以上. 从客

观指标上可以间接证明,本文算法能够保持一定的图像细节和特征,使得在进行去雾时能够很好地恢复出无雾图像,去雾图像与原始图像十分接近.

4.3.3 图像去雾主观效果

鉴于篇幅,本文只列出部分图像去雾效果,如图 16、图 17. 网络结构、参数设置、训练集不变,将有雾图像作为输入,无雾图像作为 ground truth.

图 16(b)的天空能够恢复出蓝色,云朵在去雾后十分明显;图 16(d)中山的边缘去雾完全. 图 17(d)中能够很好地恢复出街道和车辆、路牌等细节. 该结果证明本文算法有很好的迁移性. 不仅能应用于雾霾场景的生成,也可以调换训练集,训练出图像去雾的模型,且去雾效果明显,尤其在景物边缘,没有雾的残留,有很好的处理效果.

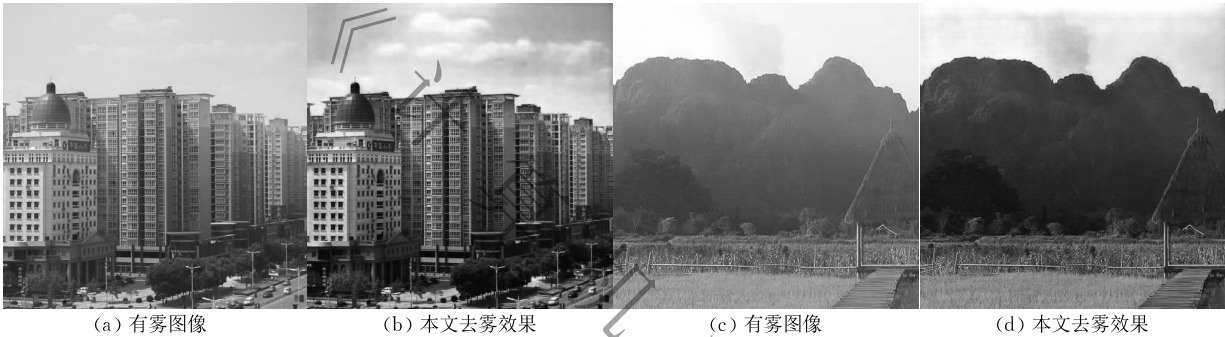


图 16 本文去雾结果 1

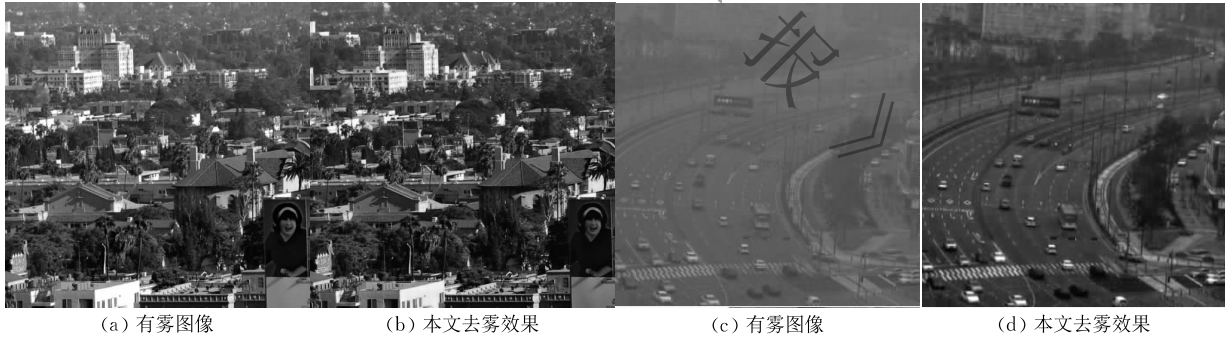


图 17 本文去雾结果 2

5 本文总结

本文设计了基于生成对抗网络 GAN 的雾霾图像转换算法,可以同时实现图像去雾和加雾. 对于加雾效果的实现,采用无雾图像作为输入,并训练 GAN 网络,生成雾天图像. 同时,本文网络结构的设计分别在生成器和判别器上进行,分别采用改进 U-net 网络和多层下采样卷积操作设计,而所有层的参数可以从训练中自动学习. 并且能够通过调整

判别器网络深度,及各项参数如学习率、损失函数、迭代次数等来获得较好的输出图像,且加雾图像细节保持较好,块效应基本消除. 基准图像的实验表明,本文算法能够提供较好的雾霾合成效果,为深度学习的去雾算法提供稳定的图像匹配训练对.

参 考 文 献

[1] Peng Ya-Li, Zhang Lu, Zhang Yu, et al. Deep deconvolution neural network for image super resolution. Journal of Software, 2018, 29(4): 926-934(in Chinese)

- (彭亚丽, 张鲁, 张钰等. 基于深度反卷积神经网络的图像超分辨率算法. 软件学报, 2018, 29(4): 926-934)
- [2] Xiao Jin-Sheng, Tian Hong, Zou Wen-Tao, et al. Stereo matching based on convolutional neural network. *Acta Optica Sinica*, 2018, 38(8): 0815017(7)(in Chinese)
(肖进胜, 田红, 邹文涛等. 基于深度卷积神经网络的双目立体视觉匹配算法. 光学学报, 2018, 38(8): 0815017(7))
- [3] Liu D, Klette R. Fog effect for photography using stereo vision. *Visual Computer*, 2016, 32(1): 1-11
- [4] Liu Jie-Ping, Huang Bing-Kun, Wei Gang. A fast effective single image dehazing algorithm. *Acta Electronica Sinica*, 2017, 45(8): 1896-1901(in Chinese)
(刘杰平, 黄炳坤, 韦岗. 一种快速的单幅图像去雾算法. 电子学报, 2017, 45(8): 1896-1901)
- [5] Zhao F, Zeng M, Jiang B, et al. Render synthetic fog into interior and exterior photographs//Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry. Hong Kong, China, 2013: 157-166
- [6] Hertzmann A, Jacobs C E, Oliver N, et al. Image analogies//Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA, 2001: 327-340
- [7] Gatys L A, Ecker A S, Bethge M. Image style transfer using convolutional neural networks//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA, 2016: 2414-2423
- [8] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets//Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada, 2014: 2672-2680
- [9] Isola P, Zhu J, Zhou T, Efros A A. Image-to-image translation with conditional adversarial networks//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii, USA, 2017: 5967-5976
- [10] Tian Wei, Lv Jingyuan. Synthesis of image patterns based on cooperative training ConvNet//Proceedings of the 2017 IEEE International Conference on Information and Automation (ICIA). Macau, China, 2017: 5427-5432
- [11] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks//Proceedings of the 4th International Conference on Learning Representations (ICLR). San Juan, Puerto Rico, 2016
- [12] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation//Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham, Germany, 2015: 234-241
- [13] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks//Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy, 2017: 2242-2251
- [14] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift//Proceedings of the 32nd International Conference on Machine Learning. Lille, France, 2015: 448-456
- [15] He K, Sun J, Tang X. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(12): 2341-2353
- [16] Tarel Jean-Philippe, et al. Vision enhancement in homogeneous and heterogeneous fog. *IEEE Intelligent Transportation Systems Magazine*, 2012, 4(2): 6-20
- [17] Wang Chao, Zheng Hai-Yong, Yu Zhi-Bin, et al. Discriminative region proposal adversarial networks for high-quality image-to-image translation//Proceedings of the European Conference on Computer Vision. Munich, Germany, 2018: 770-785
- [18] Choi L K, You J, Bovik A C. Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Transactions on Image Processing*, 2015, 24(11): 3888-3901



XIAO Jin-Sheng, Ph. D., associate professor. His research interests include computer vision, image and video processing.

SHEN Meng-Yao, M. S. candidate. Her current research interest focuses on computer vision.

LEI Jun-Feng, Ph. D., associate professor. His current research interests include image and video processing, computer vision.

XIONG Wen-Xin, M. S. candidate. Her current research interest focuses on digital image processing.

JIAO Chen-Kun, M. S. His current research interest focuses on computer vision.

Background

Generative adversarial networks cannot learn the mapping between image pixels resulting in the uncontrollable generated images. This paper proposes a new method based on GAN to learn the mapping between input and output image. The

generator and discriminator are improved to train the mapping between pixels in hazy images and haze-free images. Taking synthesis of hazy image as an example, we propose two networks: a hourglass-shaped fast link generator

network using the fog-free image as the input and a funnel-shaped global convolutional judger network to test the results of smog image synthesis. The modified loss function is to make the training model more accurate. The effect of our algorithm is better than state-of-the-art methods on both synthetic and real-world images qualitatively and quantitatively. There is no color distortion in our haze removal task.

The generated network is divided into two parts, encoding and decoding, and the underlying texture information of the image is preserved by adding corresponding convolution layers. The funnel-shaped global convolutional judger network is designed to test the results of smog image synthesis, it's

not only to fool the discriminator but also to be near the ground truth Output. The loss function of network is modified to calculate the sum of GAN loss and ABS loss, better results can be gotten. With all of these improvements, the model can learn the mapping accurately and avoid deviation and distortion well. The experimental results show the practicability and the advancement of the proposed algorithm.

This work is supported by the National Key R&D Program of China under Grant No. 2016YFB0502602 and the National Natural Science Foundation of China under Grant No. 61471272.

