



## Unsupervised Single Image Dehazing Using Dark Channel Prior Loss

---

Xiangyu Li

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

June 11, 2023



**北京航空航天大学**  
BEIHANG UNIVERSITY

## 基于暗通道先验损失的无监督单图像去雾

20373380 李翔宇

2021 年 2 月

# 基于暗通道先验损失的无监督单图像去雾

## 摘要

图像去雾(Haze Removal) 是图像处理领域越发热门的研究方向之一，其任务是将受雾霾影响的模糊图片处理为晴朗天气下的清晰图片。经典的基于图像先验的去雾方法，通常将去雾问题表述为能量最小化任务，需要手动制作输入和输出示例来提供信息。而随着图像处理技术的发展，出现了众多基于大量数据学习的技术，但它们对示例数量的需求很大。由于很难获取相同场景在晴朗和雾霾条件下的图像，并且现有的室外图像数据集的深度信息非常不精确，导致对室外图像的处理尤为困难。为了解决该问题，本文提出了一种无监督的、基于上下文聚合网络(Context Aggregation Network, CAN)的架构，该架构使用暗通道先验(Dark Channel Prior, DCP)能量函数来优化网络权重，不依赖于原始模糊图像之外的任何数据即可得到对应的清晰图像。

**关键词：** 图像去雾，无监督，暗通道先验

# 目录

目录 .....	3
1 研究背景 .....	4
1.1 基于先验的方法 .....	4
1.2 数据驱动方法 .....	4
2 问题描述 .....	5
3 解决方法 .....	5
3.1 雾霾模型(Haze Model) .....	5
3.2 暗通道先验(Dark Channel Prior) .....	5
3.3 软抠图(Soft Matting) .....	6
3.4 损失函数的实现(Implementation as a Loss Function) .....	6
3.5 计算场景辐射(Computing the Scene Radiance) .....	7
3.6 架构(Architecture) .....	7
4 实验介绍 .....	8
实验方法 .....	8
实验环境 .....	8
5 实验结果与分析 .....	9
定量分析 .....	9
定性分析 .....	10
总结 .....	12
6 参考文献 .....	12

# 1 研究背景

图像去雾任务大多采用雾霾模型来进行求解。基于求解方法的不同，现有的图像去雾方法主要分为两种：基于先验的方法、数据驱动的方法。

## 1.1 基于先验的方法

对图像去雾的早期尝试需要用到同一场景在不同天气条件下<sup>[1]</sup>，或使用不同滤镜<sup>[2]</sup>得到的多幅图像。例如，Kopf等人提出了一种利用地形，城市模型等信息对室外图像进行去雾的方法，是一种基于先验知识的方法<sup>[3]</sup>。

常见的基于先验的方法有：使用 Markov Random Field (MRF) 框架最大化图像中每个斑块的局部对比度<sup>[4]</sup>、利用透射和阴影的相关性较小等先验知识来解决场景反射率和空气光之间的不一致问题<sup>[5]</sup>、利用暗通道先验 (DCP，即在清晰图像中，图像块中的最暗像素接近于零，天空除外) 进行处理<sup>[6]</sup>、“颜色线”假设 (小图像块中的像素在 RGB 空间中具有一维分布的特征)<sup>[7]</sup>等。

虽然基于先验的方法展示了精细的图像细节，但它们通常具有饱和度和对比度的增加、颜色不真实以及难以处理天空区域等缺陷。在某种程度上，这是因为假设不适用于所有的模糊图像。此外，每个图像都需要一个单独的非平凡的优化和解决方案，而这对于实时应用来说是一大障碍。

## 1.2 数据驱动方法

现有的数据驱动方法往往需要一对模糊和清晰图像的数据集，如使用颜色衰减先验 (Color Attenuation Prior, CAP)<sup>[8]</sup>的方案、使用神经网络进行单图像去模糊 (MSCNN)<sup>[9]</sup>、AOD-Net<sup>[10]</sup>、条件 GAN 架构<sup>[11]</sup>等。但它们大多数都是在有监督的条件下进行训练的，依赖于合成数据集。

## 2 问题描述

尽管基于先验知识和数据驱动的方法可以获得较好的去雾效果，但仍存在问题：

1) 基于先验知识进行去雾的方法仅利用深度学习或先验知识估计出模型参数，本质上并没有对单个图像进行实时分析和去雾，导致迁移性差，光晕伪影等问题<sup>[3-8]</sup>。

2) 现有的数据驱动方法<sup>[9-11]</sup>与基于先验知识的方法都需要去雾前后的对比图像数据集来进行监督训练，但现有室外图像集的深度信息非常不精确<sup>[12,13]</sup>，室内数据集则存在域转移问题，难以实际应用。

## 3 解决方法

### 3.1 雾霾模型(Haze Model)

常用的雾霾模型<sup>[14]</sup>如式(1)：

$$I(x) = t(x)J(x) + (1 - t(x))A, t(x) = e^{-\beta d(x)} \quad (1)$$

其中， $I(x) \in R^N$ 是雾霾图像； $J(x)$ 是没有雾霾的清晰图像； $A$ 是大气光照，分为RGB三个通道； $t(x)$ 是透射率，与视景深度 $d(x)$ 相关。

从(1)看出，去雾问题可描述为：在已知 $I(x)$ 的条件下，获得式中的 $J(x)$ 。为此，需要求解一含有 $3N$ 个方程的方程组，其中有 $4N+3$ 个未知数( $J(x), t(x), A$ )。由于方程数量小于未知数个数，因此需要额外的先验知识辅助求解。

### 3.2 暗通道先验(Dark Channel Prior)

一个图像的暗通道<sup>[6]</sup>定义为：

$$J^{dark}(x) = \min_{c \in r, g, b} ( \min_{y \in \Omega(x)} J^c(y) ) \quad (2)$$

其中， $J^c(x)$ 表示彩色图像的每个通道， $\Omega(x)$ 是中心在 $x$ 的小图像窗口。实际生活中，暗通道所对应的数值往往接近于0。

假设 $A$ 已知，且在一个小的图像窗口内，透射率 $t(x)$ 可以认为是常数。由(1)式可得：

$$\frac{I^c(x)}{A^c} = \frac{J^c(x)}{A^c} t(x) + 1 - t(x) \quad (3)$$

根据暗通道先验， $J^{dark} \rightarrow 0$ ；对(3)式两端求两次最小值运算，并引入修正因子 $w \in$

[0,1] (通常取 0.95), 使得去雾后保留正常浓度的雾, 从而得到粗透射率:

$$\tilde{t}(x) = 1 - w \min_{c \in r, g, b} \left( \min_{y \in \Omega(x)} \frac{I^c(y)}{A^c} \right) \quad (4)$$

### 3.3 软抠图(Soft Matting)

式(1)中的雾霾模型类似于[15]中提到的抠图模型:

$$I_i = a_i F_i + (1 - a_i) B_i \quad (5)$$

其输出图像是一个前景图像和背景图像的凸线性组合, 由线性组合系数  $\alpha \in [0,1]$  控制。用粗透射率  $\tilde{t}(x)$  代替该式中的  $a$  可得能量函数:

$$E(t, \tilde{t}) = t^T L t + \lambda (t - \tilde{t})^T (t - \tilde{t}) \quad (6)$$

式(6)中第一项有利于成功抠图, 第二项代表暗通道的保真度;  $\lambda$  用来确定两者之间的相关程度, 可以取  $10^{-4}$ [3];  $\delta = [i = j]$ ;  $L$  是类拉普拉斯矩阵:

$$L_{ij} = \sum_{n|(i,j) \in p_n} (\delta_{ij} - w_{ij}^n) \quad \forall i, j = 1 \dots N \quad (7)$$

$$w_{ij}^n = \frac{1}{|p_n|} \left[ 1 + (I_i - \mu_n)^T (\Sigma_n + \frac{\varepsilon}{|p_n|} U_3)^{-1} (I_j - \mu_n) \right] \quad (8)$$

其中,  $i$  和  $j$  是以像素点  $n$  为中心的小窗口  $p_n$  处的两个像素点;  $p_n$  常取  $3 \times 3$ ;  $\mu_n \in R^3, \Sigma_n \in R^{3 \times 3}$  分别是图像窗口的均值和协方差;  $U_3$  是单位矩阵。平滑系数  $\varepsilon$  取  $10^{-6}$ 。

### 3.4 损失函数的实现(Implementation as a Loss Function)

对式(6)中的第一项进行拉普拉斯矩阵分解可得:

$$E_1(t, \tilde{t}) = t^T L t = \sum_{n=1}^N \sum_{i=1}^9 \sum_{j=1}^9 w_{ij}^n (t_i - t_j)^2 \quad (9)$$

将式(9)代入式(6)并向量化可得:

$$E(t, \tilde{t}) = \sum_{n=1}^N \sum_{k=1}^K W \odot (T_\theta^I - T_\theta^J)^2 + \lambda \sum_{k=1}^K (t_\theta - \tilde{t}(I))^2 \quad (10)$$

其中,  $\odot$  表示按元素相乘;  $K = 81$ ;  $W, T_I, T_J \in R^{N \times K}$ ;  $\theta$  下标表示  $T^I, T^J, t$  是网络学

习的结果,  $T^I, T^J$  的具体形式为  $T_n^I = \begin{bmatrix} t_1 & \dots & t_9 \\ \vdots & \ddots & \vdots \\ t_1 & \dots & t_9 \end{bmatrix}, T_n^J = \begin{bmatrix} t_1 & \dots & t_1 \\ \vdots & \ddots & \vdots \\ t_9 & \dots & t_9 \end{bmatrix}$ 。

基于式(10)的损失函数训练可得到预期网络参数 $t_\theta$ ，从而将原问题转化为基于只包含雾霾图像的训练集 $\{I_m\}_{m=1}^M$ 即可进行求解：

$$\theta^* = \arg \min_{\theta} \left[ \frac{1}{M} \sum_{m=1}^M E(t_\theta, \tilde{t}(I_m)) \right] \quad (11)$$

其中， $M$ 是图像数量。损失模型的输入输出图如图 3.4.1：

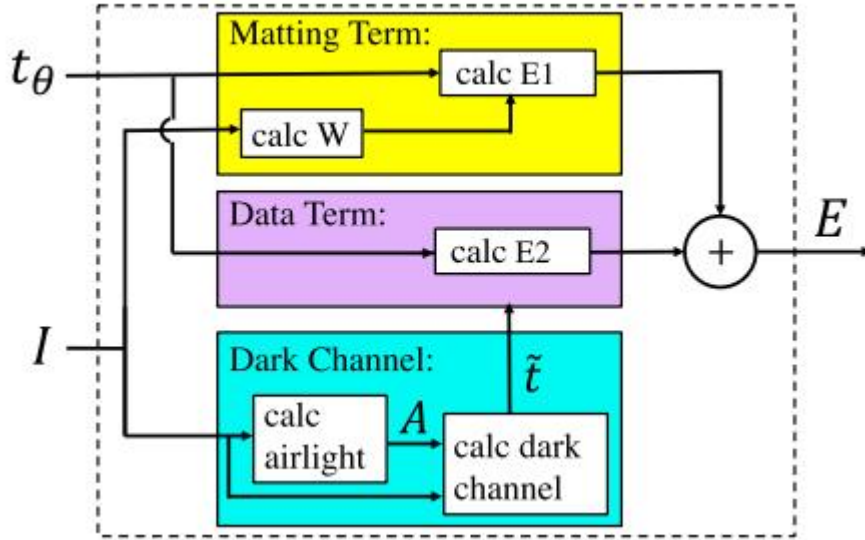


图 3.4.1

### 3.5 计算场景辐射(Computing the Scene Radiance)

网络训练完成后，新雾霾图像的传输映射函数 $t_\theta(x)$ 便能够通过向前运算得到，从而通过雾霾模型(1)来恢复场景辐射：

$$J(x) = \frac{I(x) - A}{\max(t_\theta(x), t_0)} + A \quad (12)$$

其中， $t_0$ （取 0.1<sup>[6]</sup>）用来防止分母取 0；对于大气光照 $A$ ，采用[6]中的方法：首先在雾霾图像的暗通道中选取亮度前 0.1%的点，对应到原图中再取最亮的点作为 $A$ 的估计。

### 3.6 架构(Architecture)

受到[16]的启发，本论文提出了一种名为 Dilated Residual Network 的网络，如图 3.6.1。

类似于 Context Aggregation Network (CAN)，该网络保证每个层的分辨率与输入输



出图像的分辨率一致。为了获得较为精确的预测值，本文避免了池化层和上采样，而是采用指数型扩张因子的扩张卷积来增加接收场(Receptive Field)，并在每个扩张卷积层之间增加了两个普通的卷积层，从而获得更丰富的非线性表示。

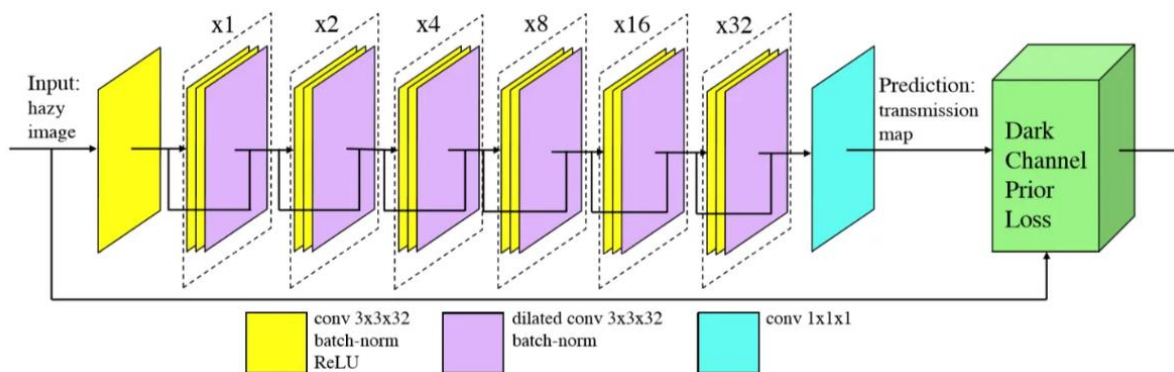


图 3.6.1

该网络由 6 个扩张残差块串联而成，每个块由两个常规卷积层和一个扩张卷积层组成。从上一个块到下一个块，扩张卷积层的扩张因子均加倍。每个卷积层的滤波尺度都是  $3 \times 3 \times 32$ 。每个常规卷积层之后都进行批归一化<sup>[17]</sup>和 ReLU<sup>[18]</sup>非线性激活函数操作，而每个空洞卷积之后只进行批归一化。最后一层是一个线性变换，将结果转换为传输图的输出维度  $1 \times 1 \times 1$ 。本文还引入了额外的 ResNet-style<sup>[19]</sup>的跳跃连接，将每个块的输入和输出简单相加，从而改善梯度流和向输出传播更细节的信息。

## 4 实验介绍

### 4.1 实验方法

论文使用 RESIDE (REalistic Single Image DEhazing) 的数据集 RTTS 和 OTS (包括 SOTS-indoor、SOTS-outdoor、HSTS 三个子集) 分别作为训练集和测试集<sup>[4]</sup>，以 PSNR (峰值信噪比, Peak Signal to Noise Ratio) 和 SSIM (结构相似性, Structural Similarity) 作为评价指标，将所提方法与已有的多种去雾方法进行对比分析，从而分析本文所提方法的优劣。

### 4.2 实验环境

实验环境的设置主要分为硬件环境、软件环境、参数设置、训练时间四部分。

1. 硬件环境: GTX Titan-X Nvidia GPU
2. 软件环境: 框架选择 TensorFlow; 优化器采用 Adam 优化器。
3. 参数设置: 模型的损失函数参数直接使用[6], [19]中建议的损失函数参数,  $\lambda$ 、 $\omega$ 、 $t_0$ 、 $\varepsilon$ 、DCP 应用区域和软抠图应用区域大小分别为  $10^{-4}$ 、0.95、0.1、 $10^{-6}$ 、 $15 \times 15$  和  $3 \times 3$ ; 优化器的参数设为批大小为 24, 初始学习率为  $l_r = 3 \cdot 10^{-4}$ , 指数衰减因子为 0.96 每 3 个 epoch; 网络权重初始化使用均值为 0, 方差为 0.1 的随机初始化。
4. 训练时间: 训练 8 小时, 获取约 30 个 epoch 后的最佳解决方案。对于户外数据, 在第 27 个 epoch 停止; 室内数据则在第 30 个 epoch 停止。

## 5 实验结果与分析

论文基于三个不同的测试集, 从定量分析与定性分析两方面评估了所提方法的性能。这些测试集是通过合成创建的, 因此包含了清晰图像和它们的有雾版本。

### 5.1 定量分析

论文将所提方法与基于先验的方法 (DCP [6]、BCCR [20]、NLD [21]) 以及基于数据驱动的方法 (CAP [8]、MSCNN [9]、DehazeNet [22]、AOD-Net [10]和 GFN [23]) 进行了比较。图 5.1.1 给出了论文所提方法与其他方法在不同测试集上的 PSNR/SSIM 指标。

	DCP [3]	BCCR [8]	NLD [7]	CAP [6]	MSCNN [12]	DehazeNet [11]	AOD-Net [10]	GFN [13]	Ours
<i>HSTS</i>	17.22/0.798	15.09/0.738	17.62/0.792	21.54/0.867	18.29/0.841	<b>24.49/0.915</b>	21.58/0.922	22.94/0.874	24.44/ <b>0.933</b>
<i>SOTS-outdoor</i>	17.56/0.822	15.49/0.781	18.07/0.802	22.30/0.914	19.56/0.863	22.72/0.858	21.34/0.924	21.49/0.838	<b>24.08/0.933</b>
<i>SOTS-indoor</i>	20.15/0.872	16.88/0.791	17.29/0.749	19.05/0.836	17.11/0.805	21.14/0.847	19.38/0.849	<b>22.32/0.880</b>	19.25/0.832

图 5.1.1 各去雾方法在各测试集中所得的 PSNR/SSIM 指标

从图 5.1.1 可以看出, 论文提出的方法在 SOTS-outdoor 测试集中所得结果的两项指标均为最高; 在 HSTS 测试集的结果中, PSNR 与最大值十分接近、SSIM 为最高; 在 SOTS-indoor 测试集中, 由于训练集为室外图像, 因此论文方法受到域转移的影响, 两项指标得分较低。

除此之外, 本文的方法在效率上继承了原本 DCP 方法的高效性。在加上 GPU 与 CPU 的运行时间后, 本文使用 SOTS-outdoor 数据集, 将自己的方法与使用 matlab 实现的软抠图 DCP (slow-DCP) [6]、引导图像滤波器 DCP (fast-DCP) [24]的方法进行了对

比。图 5.1.2 给出了论文方法与其他方法在 SOTS-outdoor 测试集上的 PSNR/SSIM 与运行时间指标。

	slow-DCP [3]	fast-DCP [35]	ours-CPU	ours-GPU
PSNR/SSIM	17.56/0.822	14.62/0.752	<b>24.08/0.933</b>	<b>24.08/0.933</b>
runtime[sec]	21.67	1.08	0.81	<b>0.28</b>

图 5.1.2 各去雾方法在 SOTS-outdoor 测试集中所得的 PSNR/SSIM 与运行时间指标

根据图 5.1.2 可以看出，本文的方法与两种根据 DCP 进行去雾的方法相比，PSNR 和 SSIM 的表现均有很大提升，而将 CPU 与 GPU 的运行时间相加后，也与 fast-DCP 的用时较接近，然而 fast-DCP 的去雾效果却是三种方法中最差的。

## 5.2 定性分析

图 5.2.1 展示了论文所提方法在 HSTS 测试集上的定性分析结果。

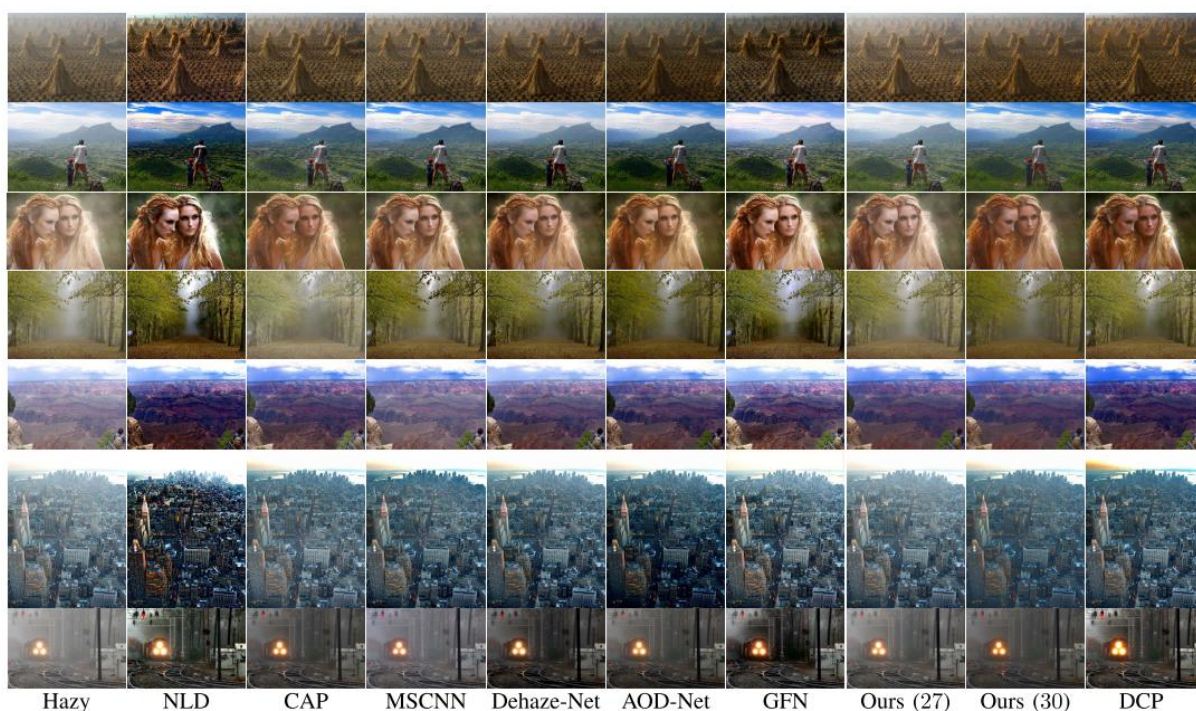


图 5.2.1 在 HSTS 测试集上的去雾效果对比

上半部分：与基于先验的方法进行比较；下半部分：与基于深度学习的方法进行比较

在图 5.2.1 的上半部分，可以看到论文的方法保持了原始图像的真实颜色，而 DCP<sup>[6]</sup>、BCCR<sup>[20]</sup>和 NLD<sup>[21]</sup>则往往会产生过饱和的天空。论文的结果与 CAP<sup>[8]</sup>产生的结果相似，但更接近地面真实图像的真实颜色。

图 5.2.1 的下半部分将论文方法与基于深度学习的方法进行了比较。在大多数图像中，论文保留了真实的对比度和颜色，而 MSCNN<sup>[9]</sup>和 GFN<sup>[23]</sup>的结果中，对比度往往过大。有时，论文所提方法会稍微改变天空的颜色，这是因为论文的训练方法是无监督的，没有清晰的天空图像供其参考。

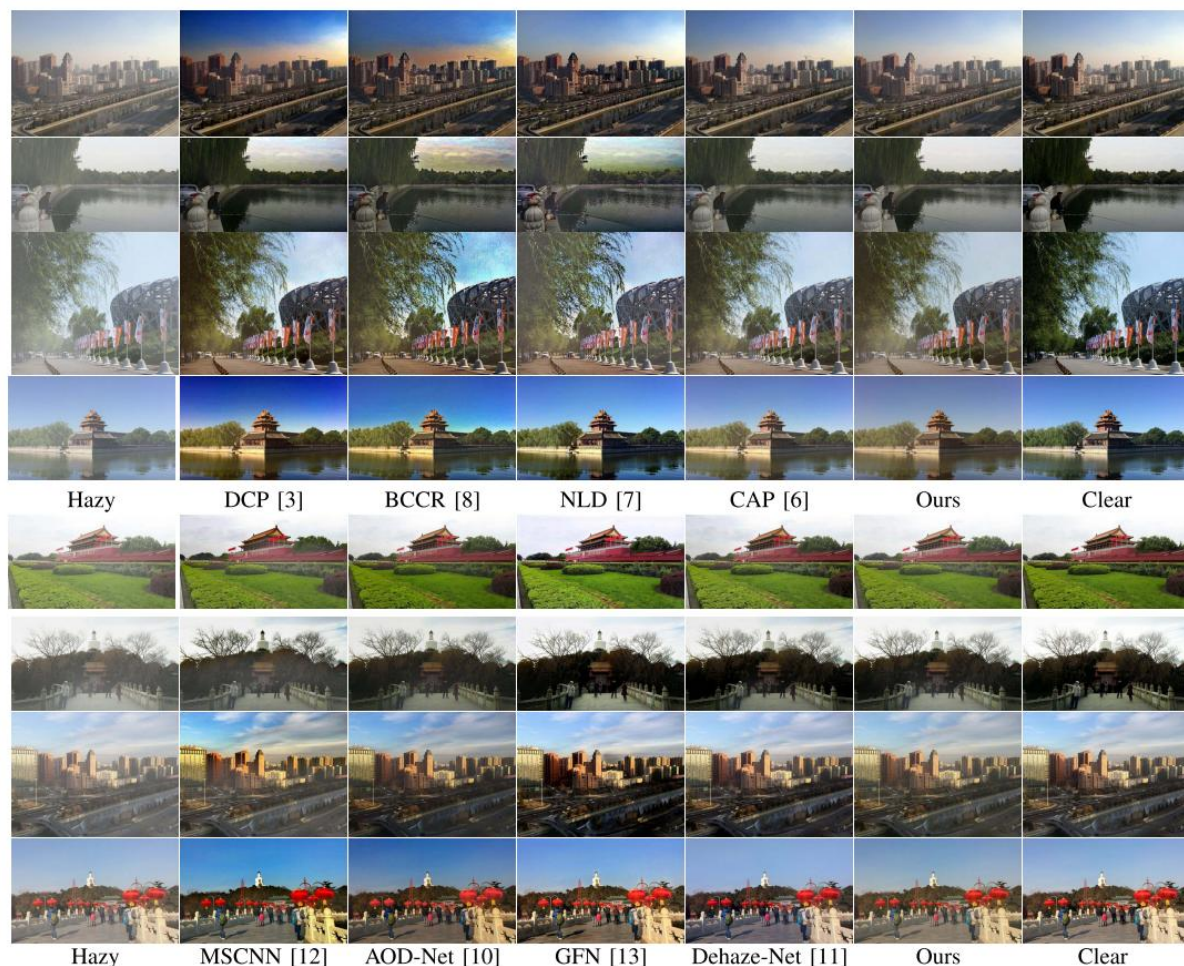


图 5.2.2 现实世界图像的去雾效果对比（括号中的数字表示训练的 epoch 数）

图 5.2.2 给出了论文的方法与基于先验和数据驱动方法的去雾效果对比。论文显示了论文的网络在 27 个 epoch（在测试中得到的最佳 epoch 数量）和 30 个 epoch（产生的图像更类似于 DCP）后的输出。可以看到，在 27 个 epoch 后，论文并没有完全去除所有的雾霾，这是因为 RESIDE 中的户外图像与真实世界中的图像相比，雾霾较轻，因此训练得到的方法去雾程度较轻。

### 5.3 总结

总的来说，论文提出了一种用于单图像去雾的对神经网络进行无监督训练的方法。论文方法在去雾效果上存在缺陷，可能导致天空颜色的失真，这是由于训练过程中没有晴朗图片作为参考对象。除此之外，论文方法在室外场景中展现了最好的去雾效果与很短的时间花费外，还避免了对合成训练集的需求。同时，论文的方法也可以结合 DCP 之外的其他能量函数，将其作为损失函数进行无监督训练。

## 6 参考文献

- [1] S. G. Narasimhan and S. K. Nayar. “Chromatic framework for vision in bad weather.” [A]. Proc. CVPR[C]. Jun. 2000:598–605.
- [2] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. “Polarization-based vision through haze” [J]. Appl. Opt. Jan. 2003:511–525.
- [3] J. Kopf et al. “Deep photo: Model-based photograph enhancement and viewing” [J]. ACM Trans. Graph. Dec. 2008, vol. 27:116.
- [4] R. T. Tan. “Visibility in bad weather from a single image” [A]. Proc. CVPR [C]. Jun. 2008:1–8.
- [5] R. Fattal. “Single image dehazing.” ACM Trans. Graph. [J]. Aug. 2008, vol. 27:72.
- [6] K. He, J. Sun, and X. Tang. “Single image haze removal using dark channel prior.” IEEE Trans. Pattern Anal. Mach. Intell. [J]. Dec. 2011, vol. 33:2341–2353.
- [7] R. Fattal, “Dehazing using color-lines.” ACM Trans. Graph [J]. Dec. 2014, vol. 34:13.
- [8] Q. Zhu, J. Mai, and L. Shao. “A fast single image haze removal algorithm using color attenuation prior.” IEEE Trans. Image Process [J]. Nov. 2015, vol. 24:3522–3533.
- [9] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang. “Single image dehazing via multi-scale convolutional neural networks.” [A]. Proc. ECCV [C]. Sep. 2016:154–169.

- [10]B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng. “Aod-net: All-in-one dehazing network.” [A]. Proc. ICCV [C]. Oct. 2017:4770–4778.
- [11]X. Yang, Z. Xu, and J. Luo. “Towards perceptual image dehazing by physics-based disentanglement and adversarial training.” [A]. in Proc. AAAI [C]. Apr. 2018:1–8.
- [12]A. Saxena, M. Sun, and A. Y. Ng. “Make3D: Learning 3D scene structure from a single still image.” IEEE Trans. Pattern Anal. Mach. Intell [J]. May 2009, vol. 31:824–840.
- [13]A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. “Vision meets robotics: The KITTI dataset.” Int. J. Robot. Res. [J]. 2013, vol. 32:1231–1237.
- [14]W. K. Middleton. “Vision through the atmosphere.” Geophysik II/Geophysics II. Springer [J]. 1957:254–287.
- [15]A. Levin, D. Lischinski, and Y. Weiss. “A closed form solution to natural image matting.” [A]. Proc. CVPR [C]. Jun. 2006:61–68.
- [16]F. Yu and V. Koltun. “Multi-scale context aggregation by dilated convolutions.” [EB/OL]. arXiv:1511.07122. Nov. 2015.
- [17]S. Ioffe and C. Szegedy. “Batch normalization: Accelerating deep network training by reducing internal covariate shift.” [EB/OL]. arXiv:1502.03167. Feb. 2015.
- [18]A. Krizhevsky, I. Sutskever, and G. E. Hinton. “Imagenet classification with deep convolutional neural networks.” [A]. Proc. NIPS [C]. 2012:1097–1105.
- [19]K. He, X. Zhang, S. Ren, and J. Sun. “Deep residual learning for image recognition.” [A]. Proc. CVPR [C]. Jun. 2016:770–778.
- [20]G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan. “Efficient image dehazing with boundary constraint and contextual regularization.” [A]. Proc. ICCV [C]. Dec. 2013:617–624.
- [21]D. Berman, T. Tali, and S. Avidan. “Non-local image dehazing.” [A]. Proc. CVPR [C]. Jun. 2016:1674–1682.
- [22]B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. “DehazeNet: An end-to-end system for single image haze removal.” IEEE Trans. Image Process. [J]. Nov. 2016, vol. 25:5187–5198.

- [23]W. Ren et al.. “Gated fusion network for single image dehazing.” [A]. Proc. CVPR [C]. Jun. 2018:3253–3261.
- [24]K. He, J. Sun, and X. Tang. “Guided image filtering.” IEEE Trans. Pattern Anal. Mach. Intell. [J]. Jun. 2013, vol. 35:1397–1409.