

DehazeNet: 用于单图像雾霾去除的端到端系统

目录

DehazeNet: 用于单图像雾霾去除的端到端系统	1
摘要	1
1. 介绍 (图像去雾的综述)	2
2. 相关工作 (回顾大气散射模型和雾霾相关特征)	4
A. 大气散射模型	4
B. 与雾霾相关的特征	6
3 拟议的 DehazeNet 结构 (DehazeNet 的细节)	7
DehazeNet 的各层设计的细节	8
B. 与传统去雾方法的联系	11
C DehazeNet 的训练	12
4 实验	14
A 模型及其性能	15
B. 过滤器编号和大小	17
C.合成图像的定量结果	18
D. 合成图像上的定量结果 (每层都作用细致说明)	19
E. 真实世界图像的定性结果	21
5.总结	21

摘要

- 单图像雾霾去除是一个具有挑战性的长期问题。现有方法使用各种约束/先验来获得似乎合理的去雾解决方案。实现雾霾去除的关键是估计输入模糊图像的介质透射图。在本文中,我们提出了一种名为 DehazeNet 的可训练的端到端系统,用于介质传输估计。**DehazeNet 将朦胧图像作为输入,并输出其介质透射图,随后用于通过大气散射模型恢复无雾图像。DehazeNet 采用基于卷积神经网络的深层架构,其层专门设计用于体现图像去雾中已建立的假设/先验。具体而言,Maxout 单位的图层用于特征提取,这可以生成几乎所有与 haze 相关的特征。我们还在 DehazeNet 中提出了一种新的非线性激活函数,称为双边整流线性单元,它能够提高恢复的无雾图像的质量。我们在提议的 DehazeNet 的组件与现有方法中使用的组件之间建立连接。基准图像的实验表明,DehazeNet 比现有方法具有更高的性能,同时保持高效和易用。**

关键词— Dehaze, image restoration, deep CNN, BReLU

1. 介绍 (图像去雾的综述)

雾霾是一种传统的大气现象，灰尘，烟雾和其他干燥的颗粒会掩盖大气的清澈。雾霾引起地面摄影领域的问题，其中可能需要密集大气的光穿透来成像远处的对象。由于光散射通过雾度颗粒的效果，这导致对象在图像中对比度损失的视觉效果。由于这些原因，在消费者摄影和计算机视觉应用中都需要去除雾度。

雾度去除是一个具有挑战性的问题，因为雾度透射取决于在不同位置变化的未知深度。各种图像增强技术已应用于从单个图像中去除雾度的问题，包括基于直方图的[1](“*Fog removal from color images using contrast limited adaptive histogram equalization*”), 基于对比度的[2](“*Contrast restoration of weather degraded images,*”)和基于饱和度的[3](“*The saturation adjustment in numerical modelling of fog*”)。另外，还提出了使用多个图像或深度信息的方法。例如，基于偏振的方法通过以不同偏振度拍摄的多个图像消除了雾度效应[4](“*Instant dehazing of images using polarization,*”)。在[2]中，基于多约束的方法应用于在不同天气条件下捕获相同场景的多个图像。基于深度的方法“*Deep photo: Model-based photograph enhancement and viewing,*”[5]需要来自用户输入或已知 3D 模型的一些深度信息。实际上，深度信息或多个模糊图像并不总是可用的。

由于使用了更好的假设和先验，单图像雾度去除最近取得了重大进展。具体地，在假设无雾图像的局部对比度远高于模糊图像的局部对比度的情况下，提出了基于马尔可夫随机场 (MRF) 的局部对比度最大化方法[6]用于去除雾度 (“*Visibility in bad weather from a single image,*”)。尽管对比度最大化方法能够获得令人印象深刻的结果，但它往往会产生过度饱和的图像。在[7](“*Single image dehazing,*”)中，提出了基于最小输入的独立分量分析 (ICA) 来消除彩色图像的雾度，但该方法耗时且不能用于处理浓雾图像。受暗物体减影技术的启发，基于无雾图像实验的经验统计发现暗通道先验 (DCP) [8](“*Single image haze removal using dark channel prior,*”), 其显示至少一个颜色通道具有一些在大部分中没有的非常低强度的像素阴霾补丁。对于暗通道先验，通过大气散射模型估计并去除雾度的厚度。然而，DCP 在天空图像中失去了去雾质量并且是计算密集型的。提出

了一些改进的算法来克服这些限制。为了提高除雾质量, Nishino 等[9](“*Bayesian defogging*,”), 用因子 MRF 对图像进行建模, 以更准确地估计高亮度的图像; 孟等人 [10](“*Efficient image dehazing with boundary constraint and contextual regularization*,”), 提出了一种有效的正则化去雾方法, 通过探索固有的边界约束来恢复无雾图像。为了提高计算效率, 使用标准中值滤波[11](“*An investigation of dehazing effects on image and video coding*,”), 中值滤波器的中值[12](“*Fast visibility restoration from a single color or gray level image*,”), 引导联合双边滤波[13](“*Physics-based fast single image fog removal*,”)和引导图像滤波器[14](“*Guided image filtering*”)来代替耗时的弱消光方法[15](“*A closed-form solution to natural image matting*,”)。近年来, 在机器学习框架中研究了与雾霾相关的先验。唐等人[16](“*Investigating haze-relevant features in a learning framework for image dehazing*,”), 将四种类型的雾霾相关的特征与随机森林算法相结合来估计介质传输。朱等人[17](“*A fast single image haze removal algorithm using color attenuation prior*,”), 创建一个线性模型, 用于在色彩衰减之前估计模糊图像的场景深度, 并用监督方法学习模型的参数。尽管取得了显着的进步, 但这些最先进的方法受到与雾霾相关的先验或启发式线索的限制 - 它们对于某些图像通常效果较差。**(单图像去雾的综述)**

从单个图像中去除雾霾是一项困难的视觉任务。相比之下, 人类的大脑可以在没有任何额外信息的情况下从自然景观中快速识别朦胧区域。人们可能会想要通过生物启发, 成功用 CNN 实现图像分类[18](“*ImageNet classification with deep convolutional neural networks*,”), 人脸识别[19](“*Robust face recognition via multimodal deep face representation*,”)和物体检测[20](“*Rich feature hierarchies for accurate object detection and semantic segmentation*,”)等高级视觉任务, 从而提出用于图像去雾的生物启发模型。实际上, 已经有一些(卷积)基于神经网络的深度学习方法最近被提出用于图像恢复/重建的低级视觉任务[21] - [23](“*Restoring an image taken through a window covered with dirt or rain*,”)。然而, 这些方法不能直接应用于单个图像雾度去除。

注意, 除了估计全局大气光亮度之外, 实现雾霾去除的关键是恢复准确的介质透射图(去雾的两个关键: 估计全局大气光亮度, 准确的介质透射率)。为此,

我们提出了 DehazeNet，这是一种基于 CNN 的可训练的端到端系统，用于介质传输估计。DehazeNet 将朦胧图像作为输入，并输出其介质透射图，随后用于通过简单的逐像素操作恢复无雾图像。（去雾的思路-操作方法）DehazeNet 的设计借鉴了图像去雾中已建立的假设/原则的思想，而其所有层的参数可以从训练模糊图像中自动学习。基准图像的实验表明，DehazeNet 比现有方法具有更高的性能，同时保持高效和易用。我们的主要贡献总结如下。

1) DehazeNet 是一个端到端系统。它直接学习和估计模糊图像与其介质透射率之间的映射关系。这是通过其深层架构的特殊设计来实现的，以体现已建立的图像去雾原理。

2) 我们在 DehazeNet 中提出了一种新的非线性激活函数，称为双边整流线性单元（BReLU）。BReLU 扩展了整流线性单元（ReLU）并证明了其在获得精确图像恢复方面的重要性。从技术上讲，BReLU 使用双边约束来减少搜索空间并改善收敛。

3) 我们在 DehazeNet 的组件与现有的去雾方法中使用的假设/先验之间建立联系，并解释 DehazeNet 通过从头到尾自动学习所有这些组件来改进这些方法。

本文的其余部分安排如下。在第二节中，我们回顾了大气散射模型和雾霾相关特征，它们提供了理解 DehazeNet 设计的背景知识。在第三节中，我们提出了拟议的 DehazeNet 的细节，并讨论它与现有方法的关系。实验在第 IV 节中介绍，然后在第 V 节中得出结论。（文章结构）

2. 相关工作（回顾大气散射模型和雾霾相关特征）

在文献中已经提出了许多图像去雾方法。在本节中，我们将简要回顾一些重要的问题，重点关注提出的大气散射模型的模型，这是模糊去雾的基本模型，以及那些对计算雾霾相关特征有用的假设。

A. 大气散射模型

为了描述模糊图像的形成，大气散射模型首先由 McCartney [25] 提出的，由 Narasimhan 和 Nayar [2] 进一步发展 [26]。大气散射模型可以正式写成：

$$I(x) = J(x)t(x) + \alpha(1 - t(x)), \quad (1)$$

其中 $I(x)$ 是观察到的模糊图像， $J(x)$ 是要恢复的真实场景， $t(x)$ 是介质透射， α 是全局大气光， x 是观察到的模糊图像中的像素。图 1 给出了一个例子。等式 (1) 中有三个未知数，并且在估计 α 和 $t(x)$ 之后可以恢复真实场景 $J(x)$ 。

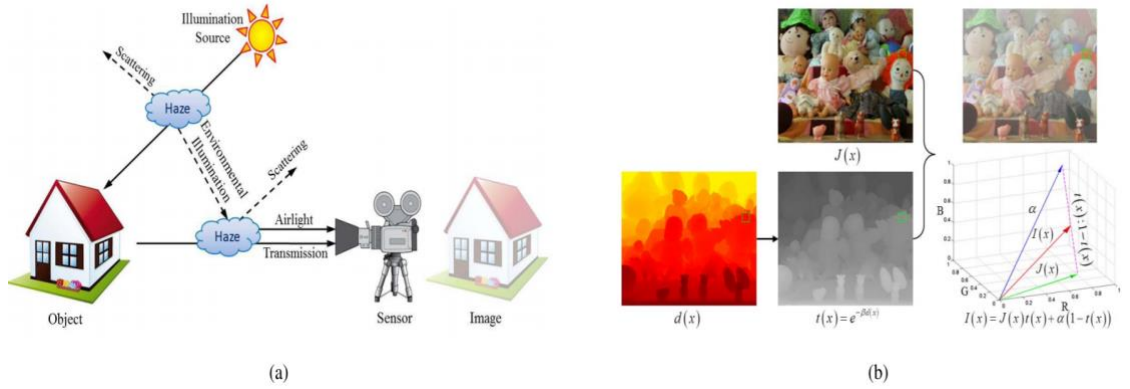


图 1. 在朦胧的天气和大气散射模型中成像。(a) 在朦胧的天气中成像的过程。由反射能量的减少引起的传输衰减 $J(x)t(x)$ 导致低亮度强度。由环境照明的散射形成的空气 $\alpha(1-t(x))$ 增强亮度和降低饱和度。(b) 大气散射模型。观察到的模糊图像 $I(x)$ 由真实场景 $J(x)$ ，介质透射 $t(x)$ 和全局大气光 α 生成。

介质透射图 $t(x)$ 描述未散射并到达相机的光部分。 $t(x)$ 定义为：

$$t(x) = e^{-\beta d(x)}, \quad (2)$$

其中 $d(x)$ 是从场景点到摄像机的距离， β 是大气的散射系数。等式 (2) 表明当 $d(x)$ 变为无穷大时， $t(x)$ 接近零。结合等式 (1) 我们有：

$$I(x) = \alpha, \quad d(x) \rightarrow \inf, \quad (3)$$

在距离视图的实际成像中， $d(x)$ 不能是无穷大，而是长距离，其给出非常低的透射率 t_0 。不是依靠等式 (3) 来获得全局大气光 α ，而是基于以下规则更稳定地估计：

$$\alpha = \max_{y \in \{x | t(x) \leq t_0\}} I(y), \quad (4)$$

上面的讨论表明，为了恢复无雾的图像（即，实现雾霾去除），估计准确的介质透射图是关键。

B. 与雾霾相关的特征

图像去雾是一个固有的长期问题。基于经验观察，现有方法提出了用于计算当前雾度相关特征的各种假设和先验知识。基于这些与雾度相关的特征，可以实现最终的除

- 1) **暗通道**：暗通道先验是基于对户外无雾图像的广泛观察。在大多数无雾度的图像中，至少一个颜色通道具有一些强度值非常低甚至接近零的像素。**暗通道[8]被定义为整个局部图像中像素通道的最小值：**

$$D(x) = \min_{y \in \Omega_r(x)} \left(\min_{c \in \{r, g, b\}} I^c(y) \right), \quad (5)$$

其中 I^c 是 I 中的RGB颜色通道上的值， $\Omega_r(x)$ 是以 x 为中心的大小为 $r \times r$ 局部图像。暗通道特征与图像中的雾度量具有高度相关性，并且用于直接估计介质透射率 $t(x) \propto 1 - D(x)$ 。

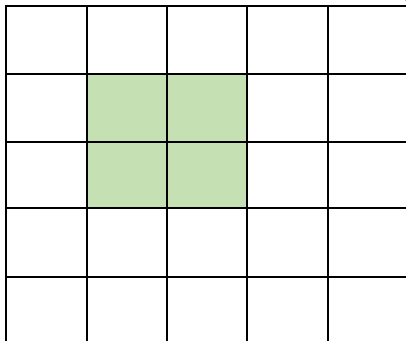
- 2) **最大对比度**：根据大气散射模型，图像的对比度通过雾度透射减少为

$$\sum_x ||\nabla I(x)|| = t \sum_x ||\nabla J(x)|| \leq \sum_x ||\nabla J(x)||。基于这种观察，局部对比度[6]$$

作为相对于中心像素的大小为 $s \times s$ 局部图像 Ω_s 中的像素强度的方差

（“*Visibility in bad weather from a single image,*”），以及大小为 $r \times r$ 的 Ω_r 区域中的局部对比度中的局部最大值被定义为：**($r \times r$ 包含很多个 $s \times s$ 块局部图像)**

$$C(x) = \max_{y \in \Omega_r(x)} \sqrt{\frac{1}{\Omega_s} \sum_{z \in \Omega_s(y)} ||I(z) - I(y)||^2}, \quad (6)$$



其中 $r=5$, 整个白色区域, $s=2$, 其中的绿色区域。

其中 $|\Omega_s|$ 是局部区域的基数。对比度特征和介质透射率 t 之间的相关性在视觉上是明显的, 因此可以通过最大化局部对比度来增强图像的可见性, 如 (6) 所示。

3) **色彩衰减**: 当场景的颜色在雾度的影响下褪色时, 图片的饱和度 $I^s(x)$ 急剧下降, 同时亮度值 $I^v(x)$ 增加, 因此产生较大的差值。根据上述 [17] 的色彩衰减先验 (“*A fast single image haze removal algorithm using color attenuation prior*,”), 利用亮度和饱和度之间的差异来估算雾度的浓度:

$$A(x) = I^v(x) - I^s(x), \quad (7)$$

其中 $I^v(x)$ 和 $I^s(x)$ 可以在 HSV 颜色空间中表示为 (亮度值) $I^v(x) =$

$$\max_{c \in \{r, g, b\}} I^c(x) \text{ 和 (饱和度) } I^s(x) = (\max_{c \in \{r, g, b\}} I^c(x) - \min_{c \in \{r, g, b\}} I^c(x)) / \max_{c \in \{r, g, b\}} I^c(x)。$$

颜色衰减特征与场景深度成正比 $d(x) \propto A(x)$, 并且容易用于传输估计。

4) **色调差异**: 原始图像 $I(x)$ 与其半逆图像之间的色调差异, $I_{si}(x) =$

$$\max_{c \in \{r, g, b\}} [I^c(x), 1 - I^c(x)], \text{ 已被用于检测雾霾。对于无雾图像, 其半反图像}$$

的三个通道中的像素值将不会全部翻转, 导致 $I_{si}(x)$ 和 $I(x)$ 之间的大的色调变化。在 [27] 中 (“*A fast semiinverse approach to detect and remove the haze from a single image*,”), 定义了色调差异特征:

$$H(x) = |I_{si}^h(x) - I^h(x)|, \quad (8)$$

其中上标 “h” 表示 HSV 颜色空间中图像的色调通道。根据 (8), 介质传输 $t(x)$ 对于 $H(x)$ 反向传播。

3 拟议的 DehazeNet 结构 (**DehazeNet 的细节**)

第 II-A 部分中的大气散射模型表明, 估算介质透射图是恢复无雾图像的最重要步骤。为此, 我们提出了 DehazeNet, 这是一种可训练的端到端系统, 可以明确地学习模糊图像与其相关介质传输图之间的映射关系。在本节中, 我们将介绍 DehazeNet 的层设计, 并讨论这些设计如何与现有图像去雾方法中的想法相关联。从估计的介质传输图中获得恢复的无雾图像的最终像素操作将在第 IV 节中给出。

DehazeNet 的各层设计的细节

拟议的 **DehazeNet** 由级联卷积和池化层组成，在这些层之后采用适当的非线性激活函数。图 2 显示了 **DehazeNet** 的架构。DehazeNet 的层和非线性激活为了实现估计介质传输做了以下四个顺序操作，即特征提取，多尺度映射，局部极值和非线性回归。我们详细介绍这些设计如下。

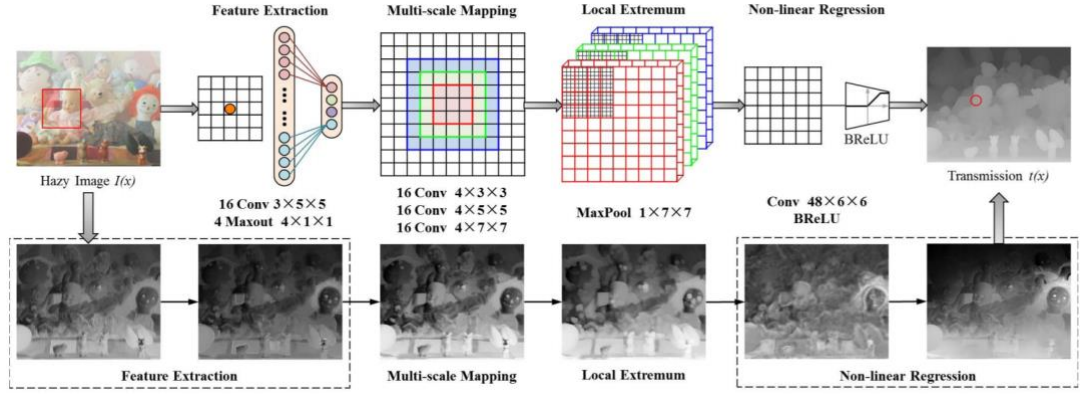


图 2. DehazeNet 的架构。DehazeNet 在概念上由四个顺序操作（特征提取，多尺度映射，局部极值和非线性回归）组成，它由 3 个卷积层，最大值池化层，Maxout 单元和 BReLU 激活函数构成。

1) **特征提取**: 为了解决图像去雾问题，现有方法提出了各种假设，并且基于这些假设，它们能够在图像中密集地提取与雾度相关的特征（例如，暗通道，色调差异和颜色衰减）。请注意，密集提取这些与雾霾相关的特征相当于将输入模糊图像与适当的滤波器进行卷积，然后进行非线性映射。受这些与雾霾相关的特征的颜色通道中的极值处理的启发，选择称为 Maxout 单元 [28] 的不寻常的激活函数作为用于降维的非线性映射。**Maxout 单位**是用于多层感知器或 CNN 的简单前馈非线性激活函数。当在 CNN 中使用时，它通过对 k 仿射特征映射进行逐像素最大化操作来生成新的特征映射。基于 Maxout 单元，我们设计 DehazeNet 的第一层如下：

$$F_1^i(x) = \max_{j \in [1, k]} g^{i,j}(x), g^{i,j} = W_1^{i,j} * I + B_1^{i,j}, \quad (9)$$

其中 $W_1 = \{W_1^{i,j}\}_{(i,j)=(1,1)}^{(n_1,k)}$ 并且 $B_1 = \{B_1^{i,j}\}_{(i,j)=(1,1)}^{(n_1,k)}$ 分别表示滤波器和偏置，

*表示卷积运算。这里，第一层有 n_1 个输出特征图。 $W_1^{i,j} \in \mathbb{R}^{3 \times f_1 \times f_1}$ 是总 $k \times n_1$ 个卷积滤波器之一，其中 3 是输入图像 $I(x)$ 中的信道数， f_1 是滤波器的大

小（详见表一）。Maxout 单位将每个 $k \times n_1$ 维向量映射到 n_1 维向量，并通过自动学习而不是现有方法中的手动方法提取与雾度相关的特征。

2) **多尺度映射**: 在[16]中, 已经证明多尺度特征对于去雾是有效的, 其在多个空间尺度上密集地计算输入图像的特征。多尺度特征提取对于**实现尺度不变性**也是有效的。例如, GoogLeNet [29]中的“Inception 架构”使用具有不同滤波器大小的**并行卷积**, 并且更好地**解决了输入图像中对齐对象的问题**, 从而在 ILSVRC14 中产生了最先进的性能[30]。在多尺度特征提取的这些成功的推动下, 我们选择在 **DehazeNet** 的第二层中使用并行卷积运算, 其中任何卷积滤波器的大小在 $3 \times 3, 5 \times 5$ 和 7×7 之间, 并且我们使用相同数量的三种过滤器。形式上, 第二层的输出写为 F 。

$$F_2^i = W_2^{[i/3],(i \setminus 3)} * F_1 + B_2^{[i/3],(i \setminus 3)}, \quad (10)$$

其中 $W_2 = \{W_2^{p,q}\}_{(p,q)=(1,1)}^{(3,n_2/3)}$ 和 $B_2 = \{B_2^{p,q}\}_{(p,q)=(1,1)}^{(3,n_2/3)}$ 包含 n_2 对参数, 分为 3 组。 **n_2 是第二层的输出维度**, $i \in [1, n_2]$ 索引输出特征映射。 $[]$ 向上取整数, \setminus 表示求余运算。

3) **局部极值**: 为了实现空间不变性, 视觉皮层中的皮层复合细胞接收来自简单细胞的响应以进行线性特征整合。Lampl 等人[31]提出复杂细胞的空间整合特性可以通过一系列池化操作来描述。根据 CNN 的经典架构[32], 在**每个像素下考虑邻域最大值以克服局部灵敏度**。另外, **局部极值是根据介质传输局部连续的假设, 并且通常克服传输估计的噪声**。因此, 我们在 **DehazeNet** 的第三层使用局部极值运算。

$$F_3^i(x) = \max_{y \in \Omega(x)} F_2^i(y), \quad (11)$$

其中 $\Omega(x)$ 是以 x 为中心的 $f_3 \times f_3$ 区域, 第三层的输出维数 $n_3 = n_2$ 。与通常降低特征图分辨率的 CNN 中的最大池相比, 此处的局部极值运算密集地应用于每个特征映射像素, 并且能够保持分辨率以用于图像恢复。

4) **非线性回归**: 深度网络中非线性激活函数的标准选择包括 Sigmoid [33]和整流线性单元 (ReLU)。前者(sigmoid)更容易发生梯度消失, 这可能导致网络训练中收敛缓慢或局部最优。为了克服梯度消失的问题, 提出了 ReLU [34], 它提供了稀疏表示。但是, **ReLU 是针对分类问题而设计的, 并不完全**

适用于图像恢复等回归问题。特别是，ReLU 仅在值小于零时才禁止值。它可能会导致响应溢出，尤其是在最后一层，因为对于图像恢复，最后一层的输出值应该在较小范围内既有下限也有上限。为此，我们提出了一种双边整流线性单元（BReLU）激活功能，如图 3 所示，以克服这一限制。受 Sigmoid 和 ReLU 的启发，BReLU 作为一种新颖的线性单元保持了双边约束和局部线性。基于所提出的 BReLU，第四层的特征图被定义为：

$$F_4 = \min(t_{\max}, \max(t_{\min}, W_4 * F_3 + B_4)), \quad (12)$$

其中 $W_4 = \{W_4\}$ 包含一个大小为 $n_3 \times f_4 \times f_4$ 的滤波器， $B_4 = \{B_4\}$ 包含一个偏差， $t_{\min, \max}$ 是 BReLU 的边际值（本文中 $t_{\min} = 0$ 和 $t_{\max} = 1$ ）。根据 (12)，该激活函数的梯度可以表示为：

$$\frac{\partial F_4(x)}{\partial F_3} = \begin{cases} \frac{\partial F_4(x)}{\partial F_3}, & t_{\min} \leq F_4(x) < t_{\max} \\ 0, & \text{otherwise} \end{cases}, \quad (13)$$

上述四层级联在一起以形成基于 CNN 的可训练端到端系统，其中与卷积层相关联的滤波器和偏置是需要学习的网络参数。我们注意到，这些层的设计可以与现有图像去雾方法的专业知识相关联，我们将在后续部分中指出。

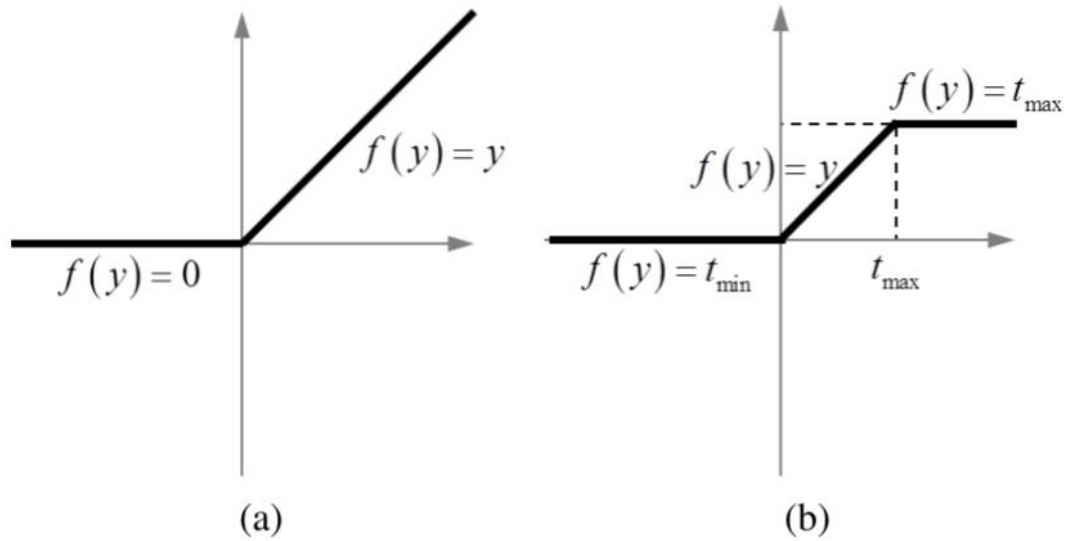


图 3. 整流线性单元（ReLU）和双边整流线性单元（BReLU）。(a) ReLU, (b) BReLU。

B. 与传统去雾方法的联系

DehazeNet 中的第一层特征 F1 专为雾霾相关特征提取而设计。以暗通道特征[8]为例。如果权重 **W1** 是相反的滤波器（在一个通道的中心处具有值-1 的稀疏矩阵，如图 4 中）并且 **B1** 是单位偏差，则特征图的最大输出等于颜色通道的最小值，类似于暗通道[8]（见公式（5））。同样，当权重是圆形滤波器时，如图 4 所示，**F1** 类似于最大对比度[6]（见公式（6））；当 **W1** 包括全通滤波器和相反滤波器时，**F1** 类似于最大和最小特征映射，它们是从 RGB 到 HSV 的颜色空间转换的操作，然后是颜色衰减[17]（见公式（7））和色调差异[27]（见等式（8））特征的提取。总之，在图 4 所示的滤波器学习成功后，第二节 B 中提到的雾度相关特征可以从 DehazeNet 的第一层中提取出来。另一方面，**Maxout** 激活函数可以被认为是对任意凸函数的分段线性近似。在本文中，我们选择四个特征图（ $k=4$ ）的最大值来逼近任意凸函数，如图 4 所示。

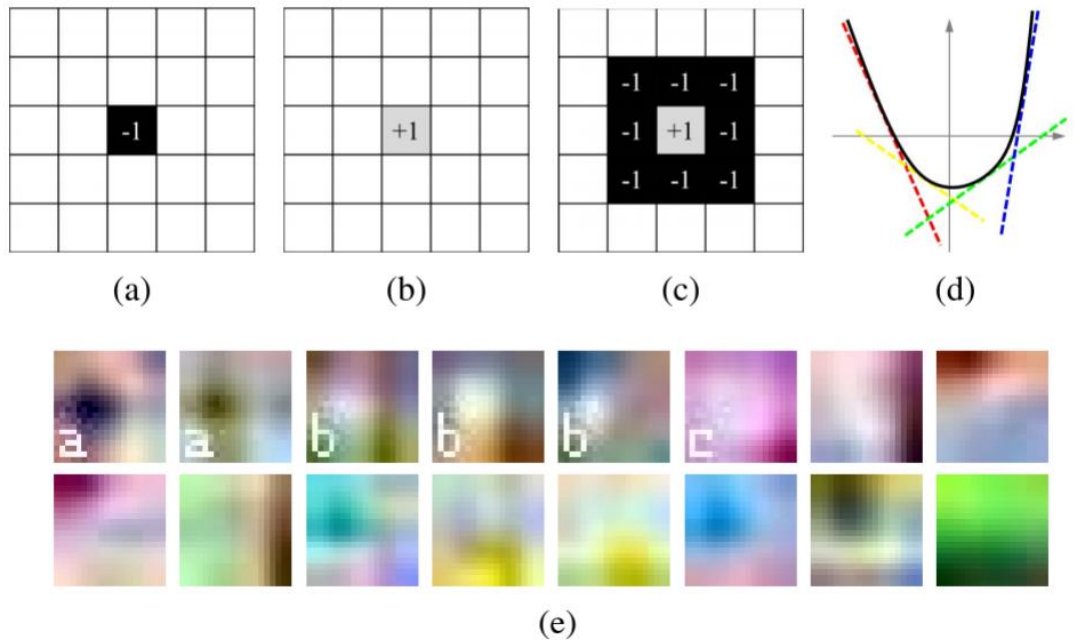


图 4. 在第一层操作 F1 中过滤器的权重和 Maxout 单位。(a) 相反的过滤器。(b) 全通滤波器。(c) 圆形过滤器。(d) Maxout。(e) 从 DehazeNet 学到的实际内核。

图像中的白色对象类似于通常具有高亮度值和低饱和度值的重雾状场景。因此，几乎所有的雾度估计模型都倾向于认为白色场景对象是远距离的，导致对介质传输的不准确估计。基于场景深度局部连续的假设，局部极值滤波器通常可以

克服这个问题[6], [8], [17]。在 DehazeNet 中, 第三层操作的局部最大滤波器消除了局部估计误差。因此, 当介质传输 $t(x)$ 接近零时, 直接衰减项 $J(x)t(x)$ 可以非常接近零。直接恢复的高亮度的场景 $J(x)$ 易于产生噪声。在 DehazeNet 中, 我们建议用 BReLU 限制 t_{min} 和 t_{max} 之间的介质传输值, 从而缓解噪声问题。请注意, BReLU 等同于传统方法中使用的边界约束[8], [17]。

C DehazeNet 的训练

1) **训练数据**: 收集大量标记数据用于训练深度模型通常是昂贵的[18]。对于 DehazeNet 的训练, 由于成对的有雾和无雾的自然场景图片 (或成对的模糊图与其相关的介质传输图) 不能大量可用, 因此更加困难。相反, 我们采用**基于物理雾霾形成模型**的综合训练数据[16]。

更具体地说, 我们基于两个假设[16]合成了朦胧和无朦胧图像的训练对**数据**: 首先, 图像内容独立于介质传输 (相同的图像内容可以出现在任何场景深度); 第二, 介质透射率是局部连续的 (图像中的局部图像像素倾向于具有相似的深度)。这些假设表明我们可以假设单个图像块的任意介质传输率。给定无雾图像 $J^P(x)$, 大气光 α 和随机介质透射率 $t \in (0,1)$, 朦胧补片合成为 $I^P(x) = J^P(x)t + \alpha(1 - t)$ 。为了减少变量学习的不确定性, 将大气光 α 设定为 1。(训练集的合成)

在这项工作中, 我们从互联网上收集无雾图像, 并从中随机抽样大小为 16×16 的图像块。与[16]不同, 这些无阴影的图像不仅包括捕捉人们日常生活的图像, 还包括那些自然和城市景观, 因为我们相信这些各种各样的培训样本可以学习到 DehazeNet 的过滤器中。图 5 显示了我们收集的无雾图像的例子。



图 5. 从互联网收集的无雾训练图像示例。

2) **训练方法**: 在 DehazeNet 中, 监督学习需要 RGB 值与介质传输之间的映射关系 F 。网络参数 $\theta = \{W1, W2, W4, B1, B2, B4\}$ 是通过最小化训练图片 $I^p(x)$ 和相应的实际介质传输 t 之间的损失函数来实现的(从公式(14)看出, 实际是最小化实际介质传输率和通过 DehazeNet 得到的介质传输率的损失函数实现的)。给定一组模糊图像及其相应的介质传输率, 其中如上所述从无模糊图像合成模糊图像, 我们使用均方误差 (MSE) 作为损失函数:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|F(I_i^p; \theta) - t_i\|^2, \quad (14)$$

其中, $F(I_i^p; \theta)$ 通过 DehazeNet 得到的介质传输率, t_i 是合成有雾的图像时的介质传输率。

随机梯度下降 (SGD) 用于训练 DehazeNet。我们使用 Caffe 包[35]实现我们的模型。我们提出的 DehazeNet (如图 2 所示) 的详细配置和参数设置总结在表 I 中, 其中包括 3 个卷积层和 1 个最大池层, 在第一次和最后一次卷积操作之后分别使用 Maxout 和 BReLU 激活。

表 1. DehazeNet 模型的体系结构

Formulation	Type	Input Size	Num n	Filter $f \times f$	Pad
Feature Extraction	Conv Maxout	$3 \times 16 \times 16$	16	5×5	0
		$16 \times 12 \times 12$	4	–	0
Multi-scale Mapping	Conv	$4 \times 12 \times 12$	16	3×3	1
			16	5×5	2
			16	7×7	3
Local Extremum	Maxpool	$48 \times 12 \times 12$	–	7×7	0
Non-linear Regression	Conv	$48 \times 6 \times 6$	1	6×6	0
	BReLU	1×1	1	–	0

4 实验

为了验证 DehazeNet 的架构，我们分析其收敛性并将其与现有技术方法进行比较，包括 FVR [12]，DCP [8]，BCCR [10]，ATM [36]，RF [16]，BPNN[37]和 CAP [17]。

关于训练数据，从因特网收集的图像中随机抽样 10,000 个无雾霾图像。对于每个图像，我们统一采样 10 个随机介质传输率 $t \in (0,1)$ 以生成 10 个朦胧图像。因此，为 DehazeNet 训练生成了总共 100,000 个合成图像（训练集）。在 DehazeNet 中，每层的滤波器权重通过从高斯分布（均值 $\mu = 0$ 和标准偏差 $\sigma = 0.001$ ）随机绘制来初始化，并且偏差设置为 0。学习率每 100,000 次迭代从 0.005 减少到 $3.125e-4$ 。基于上述参数，DehazeNet 在具有 Nvidia GeForce GTX 780 GPU 的 PC 上进行了训练（在批量大小为 128 的 500,000 次迭代中）。

基于 DehazeNet 和大气散射模型的介质传输估计，无雾图像作为传统方法得以恢复。由于第三层中的局部极值，阻塞伪像出现在从 DehazeNet 获得的传输图中。为了细化介质传输率，引导图像滤波[14]用于平滑图像。参考等式（4），在介质传输率中选择 0.1% 强度的值作为 $t(x)$ ，记为 t_0 ，并且我们选择大气亮度为 $\alpha = \max_{y \in \{x | t(x) \leq t_0\}} I(y)$ 的相关有雾的图片中的最高强度的像素。给定介质透射率 $t(x)$ 和大气光度 α ，容易恢复无雾度图像 $J(x)$ 。为了方便起见，等式（1）重写如下：

$$J(x) = \frac{I(x) - \alpha(1 - t(x))}{t(x)}, \quad (15)$$

尽管 DehazeNet 基于 CNN，但减轻的网络可以有效地保证实时性能，并且

可以在没有 GPU 的情况下运行。整个除雾框架仅在 CPU (Intel i7 3770, 3.4GHz) 中使用 MATLAB 2014A 进行测试, 处理 640×480 图像约花 1.5 秒。

A 模型及其性能

在 DehazeNet 中, 有两个重要的层, 其特殊设计用于介质传输估计——特征提取 F1 和非线性回归 F4。为了证明 DehazeNet 的有效性, 两个具有相同数量的 3 层的传统 CNN (SRCNN [38] 和 CNN-L [22]) 被视为基线模型。DehazeNet, SRCNN 和 CNN-L 的参数数量分别为 8,240, 18,400 和 67,552。

1) 特征提取 F1 中的 Maxout 单位: F1 中的激活单位是非线性降维函数, 以近似传统的方法提取的雾度相关特征。在图像处理领域, **低维映射是发现主要属性和减少模式噪声的核心过程**。例如, PCA [39] 和 LDA [40] 作为经典的线性低维映射被广泛用于计算机视觉和数据挖掘。在 [22] 中, **将非线性稀疏低维映射的 ReLU 用于高分辨率重建**。作为一种不寻常的低维映射, **Maxout 单元最大化特征映射以发现模糊图像的先验知识**。因此, 以下实验旨在确认 Maxout 单元的有效性。根据 [22], **线性单元将 16 维向量映射为 4 维向量, 相当于应用 4 个尺寸为 $16 \times 1 \times 1$ 的滤波器**。此外, 稀疏低维映射连接 ReLU 到线性单位。

图 6 显示了具有 Maxout 单元的 DehazeNet 的训练过程, 与 ReLU 和线性单元进行了比较。我们在图 6 中观察到, **Maxout 网络的收敛速度比 ReLU 和线性单元的收敛速度快**。此外, 括号中的值表示收敛结果, 与 ReLU 和线性单位相比, Maxout 的性能提高了约 $0.30e-2$ 。原因是 **Maxout 单元提供几乎所有与雾相关的特征的等效功能**, 并且减轻了诸如 ReLU 之类的简单分段功能的缺点。

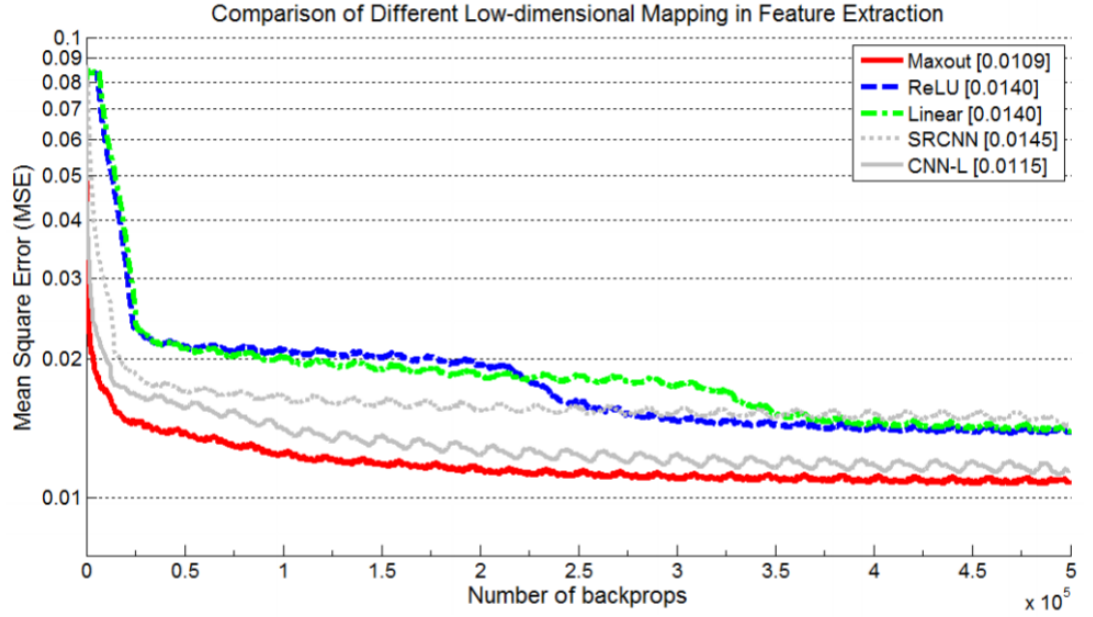


图 6. 在 F1 中使用不同的低维映射的训练过程

2) 非线性回归中的 BReLU F4: **BReLU** 是一种新颖的激活函数, 可用于图像恢复和重建。受到 ReLU 和 Sigmoid 的启发, BReLU 采用双边约束和局部线性设计。双边约束应用先验约束来减少解空间尺度; 局部线性度克服了梯度消失, 以获得更好的精度。在对比实验中, ReLU 和 Sigmoid 用于代替 BReLU 在非线性回归层中。对于 ReLU, 可以将 F4 重写为 $F_4 = \max(0, W_4 * F_3 + B_4)$, 对于 Sigmoid, 可以将其重写为 $F_4 = 1/(1 + \exp(-W_4 * F_3 - B_4))$ 。

图 7 显示了在 F4 中使用不同激活函数的训练过程。BReLU 具有比 ReLU 和 Sigmoid 更好的收敛速度, 特别是在前 50,000 次迭代期间。收敛精度表明, 与 ReLU 相比, BReLU 的性能提高了约 $0.05e-2$, 与 Sigmoid 相比提高了 $0.20e-2$ 。图 8 绘制了测试图像上预测的介质透射率与实际介质透射率的关系曲线。显然, 预测的介质透射率中心位于 BReLU 的 45 度线附近。然而, ReLU 的预测介质透射率总是高于真实介质透射率, 并且有一些预测介质透射率超过极限值 $t_{max} = 1$ 。由于 Sigmoid 函数的曲率, 预测介质透射率远离真实介质透射率, 接近 0 和 1。BReLU 测试集上的 MSE 为 $1.19e-2$, ReLU 和 Sigmoid 的 MSE 分别为 $1.28e-2$ 和 $1.46e-2$ 。

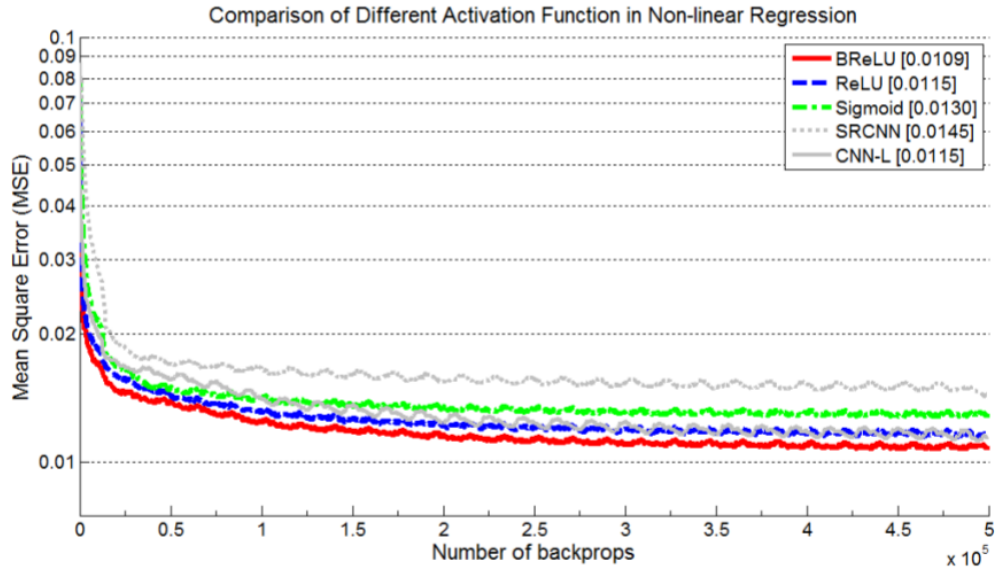


图 7. 在 F4 中使用不同激活函数的训练过程。

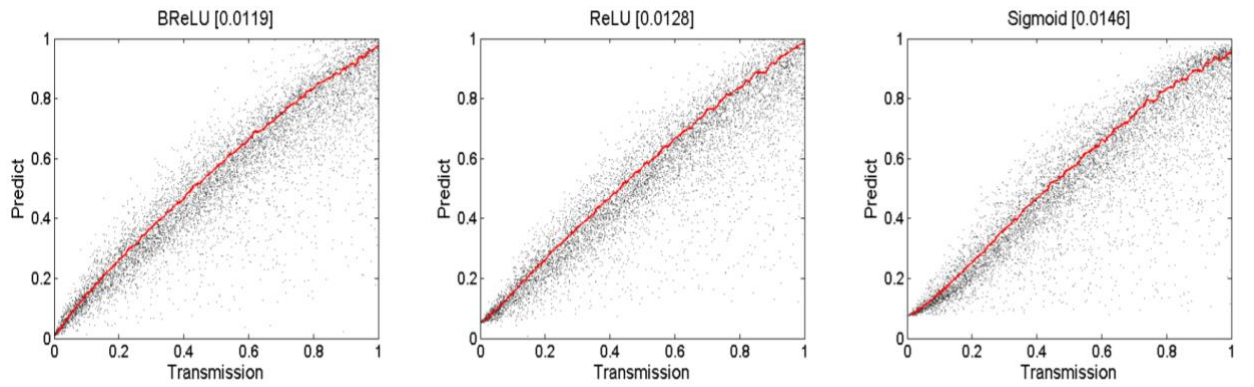


图 8. 在非线性回归 F4 中不同激活函数的预测和真实传递之间的关系图。

B. 过滤器编号和大小

为了研究性能和参数多少之间的最佳平衡，我们逐步修改 DehazeNet 的参数。根据 DehazeNet 的默认设置，进行了两个实验：（1）一个具有较大的过滤器编号，（2）另一个具有不同的过滤器大小。类似于第三节的 C.2 中，这些模型在相同的数据集上训练，并且表 II 显示出了具有相应参数设置的训练/测试 MSEs。

通常，在增加网络深度时性能会提高。同时，通过增加(每一层)滤波器的数量可以实现优越的性能。然而，如果需要快速除雾速度，则优选小型网络，其仍然可以实现比其他流行方法更好的性能。在本文中，在以下实验中采用了减轻的网络。

此外，我们还研究了不同滤波器尺寸的网络灵敏度。默认网络设置，其详

细信息如表 I 所示，表示为 5-M-7-6。我们首先分析第二层 F2 中不同滤波器尺寸的影响。表 II 表明，F2 中相当大的滤波器尺寸可以抓住更丰富的结构信息，从而产生更好的结果。在 DehazeNet 的 F2 中也采用了具有 3/5/7 滤波器尺寸的多尺度特征映射，其实现了与 7×7 滤波器的单尺度情况类似的 MSE 测试。此外，我们在第四节-D 节中证明了多尺度映射能够提高尺度稳健性。

我们进一步研究了第三层和第四层中具有不同滤波器尺寸的网络。保持相同的网络接收区域，同时调整 F3 和 F4 中的滤波器大小。结果表明，非线性回归 F4 的滤波器尺寸越大，拟合效果越好，但可能导致过拟合。F3 中的局部极值可以提高测试数据集的稳健性。因此，我们在 DehazeNet 中找到 F3 和 F4 的最佳滤波器设置为 5-M-7-6。

表 2. 在 DehazeNet 中使用不同滤波器数量或大小的结果 (×10-2)

Filter	Architecture	Train MSE	Test MSE	#Param
Number (n_1 - n_2)	4-(16×3)	1.090	1.190	8,240
	8-(32 × 3)	0.972	1.138	27,104
	16-(64 × 3)	0.902	1.112	96,704
F_2 Size (f_1 - f_2 - f_3 - f_4)	5-3-7-6	1.184	1.219	4,656
	5-5-7-6	1.133	1.225	7,728
	5-7-7-6	1.021	1.184	12,336
	5-M-7-6	1.090	1.190	8,240
F_4 Size (f_1 - f_2 - f_3 - f_4)	5-M-6-7	1.077	1.192	8,864
	5-M-7-6	1.090	1.190	8,240
	5-M-8-5	1.103	1.201	7,712

C.合成图像的定量结果

近年来，基于学习的去雾框架有三种方法。在[17]中，通过线性模型学习去雾参数，在彩色衰减先验(CAP)下，估计场景深度。反向传播神经网络(BPNN) [37]用于挖掘训练样本中颜色和深度之间的内部联系。在[16]中，随机森林(RF)用于研究雾度相关的特征，用于图像去雾恢复。所有上述方法和 DehazeNet 都采用与 RF 相同的方法进行训练。根据 RF 的测试方法，从无雾图像中随机采样 2000 个图像，其中 10 个随机介质透射率 $t \in (0,1)$ ，以生成 20,000 个朦胧图像用于测试。我们在相同的测试数据集上运行 DehazeNet 和 CAP，以测量预测介质透射率与真实介质透射率之间的均方误差(MSE)。DCP [8]是一种经典的去雾方法，用作比较基线。

表 III 显示了测试图像上的预测介质透射率和真实介质透射率之间的 MSE。DehazeNet 取得了最佳的最新成绩，即 1.19e-2;我们的方法与文献中的下一个最

新结果（RF [16]）之间的 MSE 之差为 $0.07e-2$ 。因为在 RF 中，图像的特征值被排序以打破雾度相关特征和图像内容之间的相关性。然而，DehazeNet 所关注的内容信息对于天空区域和白色物体的介质透射率估计是有用的。此外，由于线性回归模型中的一些异常值（大于 1 或小于 0），CAP [17] 在后续实验中取得了令人满意的结果，但在该实验中表现不佳。

表 3. 在合成图像数据上预测介质透射率与实际介质透射率的均方差

Methods	DCP [9]	BPNN [40]	CAP [18]	RF [17]	DehazeNet
MSE($\times 10^{-2}$)	3.18	4.37	3.32	1.26	1.19

D. 合成图像上的定量结果（每层都作用细致说明）

为了验证完整图像的有效性，DehazeNet 在具有已知深度图 $d(x)$ 的立体图像的合成模糊图像上(测试集)进行测试，并与 DCP [8]，FVR [12]，BCCR [10]，ATM [36]，CAP 2 [17] 和 RF [16] 进行比较。在 Middlebury Stereo Datasets（2001-2006）[41]-[43] 中收集了 12 对立体图像。在图 9 中，基于（1）式从无雾立体图像合成模糊图像，并且它们被 DehazeNet 恢复为无雾图像。

为了定量评估这些方法，我们根据每对无雾图像和除雾结果之间的差异使用一系列评估标准。除了广泛使用的均方误差（MSE）和结构相似性（SSIM）[44] 指数外，我们还使用了额外的评估矩阵，即峰值信噪比（PSNR）和加权峰值信噪比（WPSNR）[45]。我们将一次通过评估（OPE）定义为传统方法，我们使用标准参数运行并报告性能评估的平均度量。在表 IV 中，通过 OPE 对 DehazeNet 与六种最先进去雾方法在所有模糊图像上进行比较（使用单个散射系数 $\beta=1$ 和纯白色大气亮度 $\alpha=1$ 合成模糊图像）。令人兴奋的是，虽然 DehazeNet 通过 MSE 损失函数进行了优化，但它在其他类型的评估矩阵上也实现了最佳性能。

表 4. MSE, SSIM, PSNR 和 WPSNR 在合成图像上的平均结果（ $\beta=1$ 和 $\alpha=1$ ）

Metric	Hazy	ATM [39]	BCCR [11]	FVR [38]	DCP [9]	CAP ² [18]	RF [17]	DehazeNet
MSE	0.0481	0.0689	0.0243	0.0155	0.0172	0.0075 (0.0068)	0.0070	0.0062
SSIM	0.9936	0.9890	0.9963	0.9973	0.9981	0.9991 (0.9990)	0.9989	0.9993
PSNR	61.5835	60.8612	65.2794	66.5450	66.7392	70.0029 (70.6581)	70.0099	70.9767
WSNR	8.5958	7.8492	12.6230	13.7236	13.8508	16.9873 (17.7839)	17.1180	18.0996

去雾效果对雾度浓度的敏感，对不同散射系数 β 的性能可能变得更差或更好。因此，我们提出了一种评估来分析对散射系数 $\beta \in \{0.75, 1.0, 1.25, 1.5\}$ 的去雾鲁棒性，这被称为系数鲁棒性评估（CRE）。如表 V 所示，CAP [17] 在散射系数中等大小时（ $\beta=0.75$ ）获得了更好的性能，但是当雾度增加时，除雾性能逐渐降低。原因是 CAP 基于预测的场景深度和假定好的散射系数（ $\beta=1$ ）下估计介质透射率。在[16]中，200 棵树用于构建随机森林以进行非线性回归，并显示出更高的系数鲁棒性。但是，每个像素中随机森林的高计算量限制了它的实用性。对于 DehazeNet，介质透射率由 F1 中的非线性激活函数（Maxout）直接估算，从而对散射系数具有出色的鲁棒性。

由于雾度粒子和光源的颜色偏移，大气光不是适当的纯白色。提出了一种大

气光稳健性评估（ARE）来分析不同大气光 α 的去雾方法。虽然 DehazeNet 是通过设置 $\alpha = 1$ 产生的样本进行训练的，但它也可以在大气光的其他值上实现更强的鲁棒性。特别是，当阳光为[1.0,1.0,0.9]时，DehazeNet 的性能优于其他方法。因此，DehazeNet 也可用于去除光晕，光晕是围绕光源的明亮环，如图 10 所示。

表 5. 通过不同散射系数，图像尺度和大气光上合成图像的 MSE

Evaluation		Hazy	ATM [39]	BCCR [11]	FVR [38]	DCP [9]	CAP ² [18]	RF [17]	DehazeNet
CRE ($\beta =$)	0.75	0.0311	0.0581	0.0269	0.0122	0.0199	0.0043 (0.0042)	<u>0.0046</u>	0.0063
	1.00	0.0481	0.0689	0.0243	0.0155	0.0172	0.0077 (0.0068)	<u>0.0070</u>	0.0062
	1.25	0.0658	0.0703	0.0230	0.0219	0.0147	0.0141 (0.0121)	<u>0.0109</u>	0.0084
	1.50	0.0833	0.0683	0.0219	0.0305	0.0134	0.0231 (0.0201)	<u>0.0152</u>	0.0127
CRE Average		0.0571	0.0653	0.0254	0.0187	0.0177	0.0105 (0.0095)	<u>0.0094</u>	0.0084
ARE ($\alpha =$)	[1.0, 1.0, 1.0]	0.0481	0.0689	0.0243	0.0155	0.0172	0.0075 (0.0068)	<u>0.0070</u>	0.0062
	[0.9, 1.0, 1.0]	0.0437	0.0660	0.0266	0.0170	0.0210	0.0073 (0.0069)	0.0071	<u>0.0072</u>
	[1.0, 0.9, 1.0]	0.0435	0.0870	0.0270	0.0159	0.0200	0.0070 (0.0067)	<u>0.0073</u>	0.0074
	[1.0, 1.0, 0.9]	0.0421	0.0689	0.0239	0.0152	0.0186	<u>0.0081</u> (0.0069)	0.0083	0.0062
ARE Average		0.0443	0.0727	0.0255	0.0159	0.0192	0.0075 (0.0068)	<u>0.0074</u>	0.0067
SRE ($s =$)	0.40	0.0478	0.0450	0.0238	0.0155	0.0102	0.0137 (0.0084)	<u>0.0089</u>	0.0066
	0.60	0.0480	0.0564	0.0223	0.0154	0.0137	0.0092 (0.0071)	<u>0.0076</u>	0.0060
	0.80	0.0481	0.0619	0.0236	0.0155	0.0166	0.0086 (0.0066)	<u>0.0074</u>	0.0062
	1.00	0.0481	0.0689	0.0243	0.0155	0.0172	0.0077 (0.0068)	<u>0.0070</u>	0.0062
SRE Average		0.0480	0.0581	0.0235	0.0155	0.0144	0.0098 (0.0072)	<u>0.0077</u>	0.0062
NRE ($\sigma =$)	10	0.0484	0.0541	0.0138	0.0150	0.0133	<u>0.0065</u> (0.0070)	0.0086	0.0059
	15	0.0488	0.0439	0.0144	0.0148	0.0104	<u>0.0072</u> (0.0074)	0.0112	0.0061
	20	0.0493	–	0.0181	0.0151	0.0093	<u>0.0083</u> (0.0085)	0.0143	0.0058
	25	0.0500	–	0.0224	0.0150	<u>0.0082</u>	0.0100 (0.0092)	0.0155	0.0051
	30	0.0508	–	0.0192	0.0151	<u>0.0085</u>	0.0119 (0.0112)	0.0191	0.0049
NRE Average		0.0495	–	0.0255	0.0150	0.0100	<u>0.0088</u> (0.0087)	0.0137	0.0055

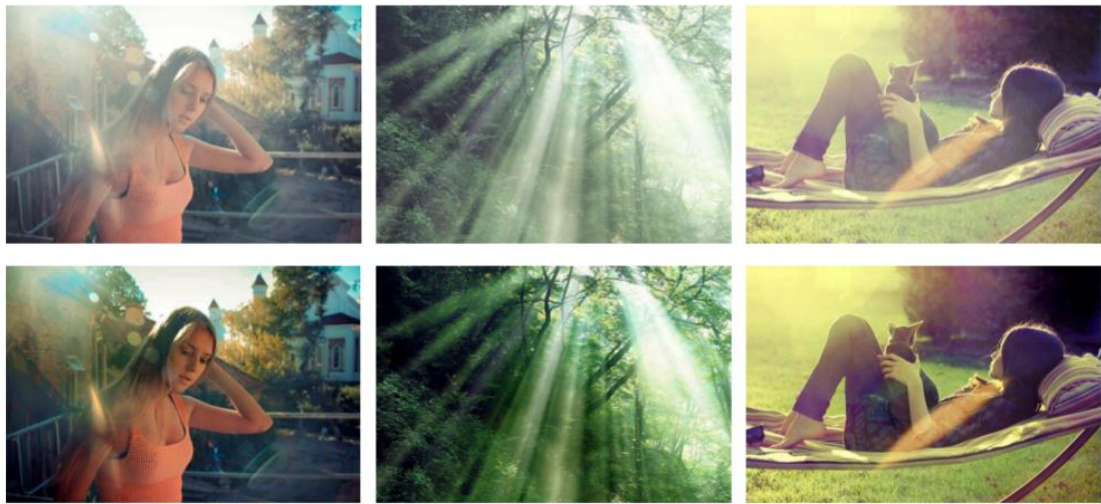


图 10. DehazeNet 的抗光晕图像增强功能。

视野变换和图像缩放经常在实际应用中发生。尺度稳健性评估（SRE）用于分析尺度变化的影响。与 OPE 中相同的现有技术方法相比，有 4 个比例系数 s 从 0.4 到 1.0 中选择，以产生 SER 的不同比例图像。在表 V 中，由于 F2 中的多尺度映射，DehazeNet 显示出对尺度变化的出色稳健性。CAP [17]，DCP [8] 和 ATM [36] 中使用的单一尺度导致不同尺度的不同预测精度。当图像缩小时，过大的处理邻域将丢失图像的细节。因此，DehazeNet 中的多尺度映射提供了

多种滤波器来合并多尺度特征，并且在所有不同尺度下都获得了最佳分数。

在大多数情况下，传感器或摄像机电路会随机产生噪声，这会带来估计误差。我们还讨论了不同程度的图像噪声对我们的方法的影响。作为基本噪声模型，具有标准偏差 $\sigma \in \{10, 15, 20, 25, 30\}$ 的加性高斯白噪声（AWG）噪声用于噪声鲁棒性评估（NRE）。受益于 F1 中的 Maxout 抑制和 F3 中的局部极值，DehazeNet 在 NRE 中的表现比其他人更强劲。RF [16] 在大多数评估中具有良好的性能但在 NRE 中失败，因为图像的特征值被分类以打破介质传输和图像内容之间的相关性，这也将放大异常值的影响。

E. 真实世界图像的定性结果

图 11 显示了 DehazeNet 恢复的去雾结果和深度图，更多结果和比较可以在 <http://caibolun.github.io/DehazeNet/> 找到。因为所有的去雾算法都可以在一般的户外图像上获得真正好的结果，所以难以在视觉上对它们进行排序。为了比较它们，本文重点关注相关研究中的 5 个已识别的具有挑战性的图像[8], [16], [17]。这些图像具有难以处理的大的白色或灰色区域，因为大多数现有的去雾算法对白色敏感。图 12 显示出了在挑战图像上与六种现有技术的去雾算法的定性比较。图 12 (a) 描绘了要去雾的模糊图像，图 12 (b-g) 分别显示了 ATM [36], BCCR [10], FVR [12], DCP [8], CAP [17] 和 RF [16] 的结果。DehazeNet 的结果如图 12 (h) 所示。

朦胧图像中的天空区域是去雾的挑战，因为云和雾是类似的自然现象与相同的大气散射模型。如前三幅图所示，大部分雾在 (b-d) 结果中被移除，并且场景和物体的细节得到很好的恢复。然而，图片结果显著地受到天空区域过度增强的影响。总的来说，这些图像的天空区域比它本来的样子更暗或者过饱和和变形。雾度通常仅存在于大气表面层中，因此天空区域几乎不需要处理。基于学习框架，CAP 和 RF 避免天空中的颜色失真，但由于是非内容回归模型（例如，第一个图像的岩石土壤和绿色平原区域），非天空区域的增强效果很差。第三张图片）。DehazeNet 似乎能够找到天空区域以保持颜色，并确保在其他区域有良好的去雾效果。原因是在 DehazeNet 的隐藏层中可以学习图像属性，它有助于天空中的去雾效果。

因为基于先验的介质透射率估计是一种统计数据，这可能对某些图像不起作用。第四和第五个数字被确定为[8]中的失败案例。当场景对象本身就类似于大气光（例如第四图中的白皙肤色和第五图中的白色大理石）时，基于先验（DCP, BCCR, FVR）的估计介质透射率是不可靠的。因为暗通道在这些物体附近具有明亮的值，并且 FVR 和 BCCR 基于 DCP，其具有过高估计介质透射率的固有问题。从回归模型中学习的 CAP 和 RF 没有过饱和，但低估了距离的雾度（参见第四张图像中的棕色头发和第五张图像中的红色柱子）。与六种算法相比，我们的结果避免了图像过饱和，并且由于 DehazeNet 的非线性回归而保留了去雾效果。

5. 总结

在本文中，我们提出了一种新的单一图像去雾的深度学习方法。受传统雾霾相关特征和去雾方法的启发，我们表明介质透射率估计可以重新构建为具有特殊设计的可训练端到端系统，其中特征提取层和非线性回归层与传统 CNN 区分开来。在第一层 F1 中，Maxout 单元被证明与先验方法类似，并且学习与雾度相关的特征更有效。在最后一层 F4 中，一种称为 BReLU 的新型激活函数代替 ReLU 或 Sigmoid，以保持图像恢复的双边约束和局部线性。凭

借这种轻量级架构，DehazeNet 比最先进的方法实现了极高的效率和出色的去雾效果。

虽然我们成功应用了 CNN 去除雾霾，但仍有一些可扩展性研究要进行。也就是说，大气光 α 不能被视为全局常数，将其与统一网络中的媒体传输一起学习。此外，我们认为大气散射模型也可以在更深的神经网络中学习，其中雾霾和无雾霾图像之间的端到端映射可以直接优化而无需介质透射率估计。我们将这个问题留待将来研究。