

针对单幅图像去雾的端到端系统

摘要：单幅图像去雾是一具有挑战性的病态问题。现有的方法使用约束或先验来获取合理的去雾方式。去雾的关键在于估计输入的含雾图像的大气透射率图。在这篇文章中，我们提出一种可训练的用于估计透射率的端到端系统，叫做去雾网络。去雾网络以含雾图像为输入，输出它的透射率图，然后用透射率图通过大气散射模型来恢复成无雾图像。去雾网络采用了基于深度结构的卷积神经网络，它的各个层都是特殊设计的，能够体现在图像去雾中已建立好的假设或先验。具体而言，最大输出单元层是作特征提取的，能够生存几乎所有的与雾有关的特征。在去雾网络中，我们也提出了一种新的非线性的激活函数，叫双边修正线性单元（BRReLU），它能够提高图像去雾的质量。我们在去雾网络和现有的去雾方法间建立起来联系。在测试图像上的实验结果表明，去雾网络在保证效率和易于使用的同时，比现有的方法表现出了更好的去雾性能。

关键词：去雾；图像复原；深度卷积神经网络；双边修正线性单元

1、介绍

雾是一种常见的大气现象。空气中的灰尘烟和其他颗粒都会降低大气的清晰度。在景物摄影领域，光在大气中的透射率对远距离物体的成像尤为重要，因此雾带来了很多问题。由于光被大气中粒子散射带来的摄影成像时物体在视觉效果上对比度的降低。因此，去雾工作在摄影领域和计算机视觉领域，都有着很大的应用前景。

去雾是一个有挑战性的问题，因为雾的透射率取决于景深，而在不同的位置，景深也不一样。各种图像增强技术已经被应用到了去雾的问题上，包括基于直方图的【1】、基于对比度的【2】和基于饱和度的【3】单张图像去雾。另外，人们也提出了使用多张图片或深度信息来去雾的方法。例如，基于【4】的偏振法，通过多张从不同偏振角度拍摄的图片来去雾。在【5】中，基于多约束的方法被应用到了在不同天气条件下拍摄的同一场景的多张图片的去雾。基于景深的【6】需要从用户的输入或已知的3D模型中得到景深信息。在实践中，利用景深或多张图片去雾并不总是可行的。

这几年由于使用了更好的假设和先验，单张图片去雾取得了很大的进步。尤其是在无雾图像的局部对比度比有雾图像更大的假设前提下，【7】提出一种基于MRF的局部对比度最大化的去雾方法。虽然对比度最大化的方法可以获得比较好的去雾性能，但这种方法容易生成过饱和的图像。在【8】中，提出了一种基于最小输入的独立成分分析（ICA）方法来除去彩色图像的雾，但是这种方法很耗时，而且不能用于浓雾图像的去雾。受到暗光物体去除技术的启发，人们基于经验数据的使用发现了暗通道先验【9】的去雾方法。这些数据来自于无雾图像的实验，实验表明在大部分的图片块中，至少有一个颜色通道有一些值很小的像素。利用暗通道先验，通过大气散射模型可以估计雾的浓度和去除雾。然而，暗通道先验的去雾方法降低了天空的成像质量，而且计算量大。人们又提出了一些改进算法来克服这些问题。为了提高去雾的质量，【10】采用阶乘的MRF方法对图像进行建模，估计得到更加准确的场景辐射。【11】提出了一种有效的正则化方法，通过探究图像内部的边界约束来恢复无雾图像。为了提高计算效率，人们使用标准中值滤波【12】、双中值滤波【13】、引导联合双边滤波【14】和引导图像滤波【15】来代替耗时的软抠图【16】步骤。在最近几年，人们在机器学习框架下进行了与雾有关的先验知识的研究。【17】利用随机森林的方法组合了四种与雾有关的特征来估计透射率。【18】基于颜色衰减先验针对有雾图像的景深估计创建了一个线性模型，并求通过有监督的方式来学习这个模型的参数。尽管去雾领域已经有了很多成就，这些先进的方法都被相

似的与雾有关的先验或启发线索限制了，这些方法处理某些图像是是很低效的。

单张图像去雾是一项困难的视觉任务。相反的，人类的大脑能够快速分辨出没有任何附加信息的自然场景中有雾的区域。参照生物卷积神经网络在高阶视觉课题如图像分类【19】、人脸识别【20】和目标检测【21】的成功应用，人们试图提出一种类似的基于生物学的模型来去雾。事实上，早就有一些用于低阶视觉任务的图像复原或重建【22】【23】【24】的基于深度学习方法的卷积神经网络被提出来了。然而，这些方式都不能被直接应用到单张图像去雾中。

注意到除了全局的大气光照的估计外，去雾的关键是估计出准确的透射率图。为了达到这个目标，我们提出了 DehazeNet，这是一个基于端到端系统的用于透射率估计的可训练的卷积神经网络。DehazeNet 以单张有雾图像作为输入，输出它的透射率图，通过简单的像素操作来复原无雾图像。DehazeNet 的设计想法源自去雾领域已建立好的假设或定理，它各个层的参数都能通过有雾图像的训练自动学习得到。测试图片上实验表明，DehazeNet 在保证效率和易于使用的同时，比现有的方法能表现出了更好的去雾性能，我们主要的贡献总结如下：

1. DehazeNet 是一个端到端系统。它可以直接学习和估计到有雾图像块和对应的透射率直接的关系。这是通过特殊设计它的深度结构来体现现有的去雾原理实现的。
2. 我们提出了一种新的非线性的激活函数，叫做双边修正线性单元 (BReLU)。BReLU 是修正线性单元 ReLU 的扩展，它对获得准确的复原图像非常重要。从技术上讲，BReLU 使用双边来限制搜索空间范围和提高收敛速度。
3. 我们将现有的去雾方法的假设或先验同 DehazeNet 的组成部分联系起来，并且解释了它通过自主学习能够获得更好的性能的原因。

这篇论文的组织结构如下。在第 2 章中，我们回顾了大气散射模型和与雾有关的特征，他们给 DehazeNet 的设计提供了背景知识。在第 3 章中，我们相信地阐述了 DehazeNet，并且讨论了它与现有方法的联系。第 4 章进行了实验，第 5 章得出结论。

2、相关工作

许多文献提出了各种各样的去雾算法。在这章中，我们简要地回顾一些重要的方法，重点关注它们提出的大气散射模型，这是图像去雾和计算与雾有关特征的的基本模型。

A. 大气散射模型

为了描述有雾图像的方程，【26】第一次提出了大气散射模型，后来被【27】【28】加以改进了。大气散射模型通常表示为

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

其中 $I(x)$ 是观测到的有雾图像， $J(x)$ 是要被复原的真实场景， $t(x)$ 是透射率， α 是全局大气光， x 是有雾图像 I 的像素坐标。图 1 给出了解释。等式(1)中有三个未知数，真实场景 $J(x)$ 可以在 α 和 $t(x)$ 估计出来后被复原。

透射率图 $t(x)$ 描述了没有被散射和到达摄像机的光的传播，它被表示为：

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

其中， $d(x)$ 表示场景点到摄像机的距离， β 是大气散射系数。等式(2)表明当距离无限远的时候，透射率接近零。根据等式(1)，有

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

实际上，距离不可能无限远，但是在相对长的距离上，透射率接近零。与其用等式(3)来获取全局大气光照，用下面的方法来估计能得到更加稳定的值：

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

以上的讨论表明为了得到一幅没有雾的图像，关键是估计出准确的透射率图。

3

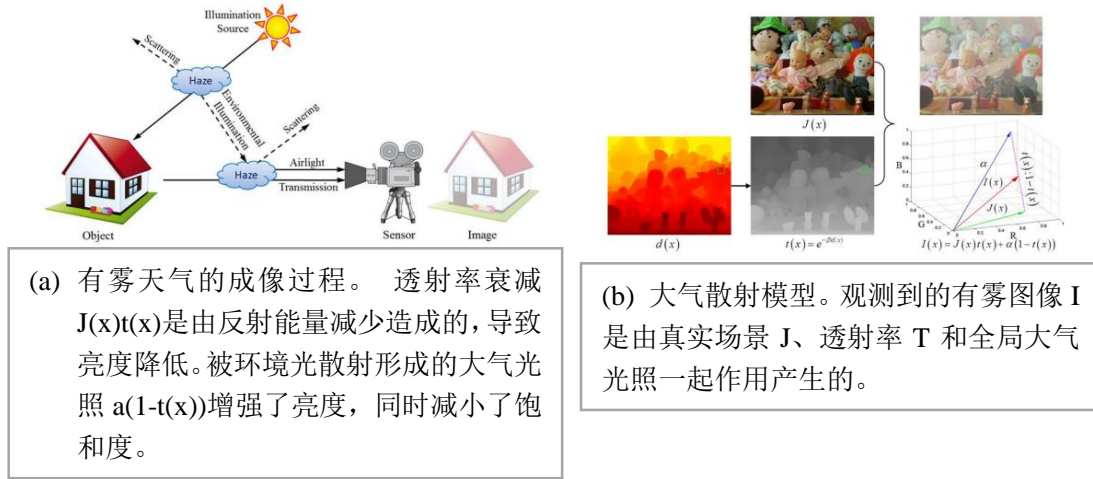


图 1. 雾天成像和大气散射模型

B. 与雾有关的特征

图像去雾本质上是一项病态问题。基于经验观察，现有的方法提出了一些假设或先验来用于计算出图像中与雾有关的特征。最后的去雾都是基于这些与雾有关的特征。

1. 暗通道：暗通道先验是基于大量的户外无雾图像的观测。大部分的无雾的图片块，至少有一个颜色通道有一些非常小或接近于 0 的像素值。暗通道【9】被定义为局部图片块的所有通道的最小像素值：

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

其中 I^c 是 I 的 RGB 颜色通道， $\Omega_r(x)$ 是以 x 为中心， $r \times r$ 为半径的局部图像块。暗通道特征和图像中雾的多少有非常大的联系，并且可以直接用来估计透射率。

2. 最大对比度：根据大气散射，图像的对比度被雾降低了，按照

$$\sum_x \|\nabla I(x)\| = t \sum_x \|\nabla J(x)\| \leq \sum_x \|\nabla J(x)\| \text{ 的方式。基于这个结论，局部对比度【7】表示为}$$

$s \times s$ 的局部图像块的像素值与中心像素值的方差，在 $r \times r$ 区域的局部对比度的最大值被定义为：

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

其中 $|\Omega_s(y)|$ 是局部邻域的基数。对比度和透射率直接的联系是显而易见的，所以可以按等式(6)通过最大化局部对比度来提高图像的可见度。

3. 颜色衰减：在雾的影响下，场景颜色衰减，此时图像块的饱和度急剧降低而亮度值增加，这样就产生了很大的差异。根据上面的颜色衰减先验【18】，亮度值和饱和值的差异被用来估计雾的浓度：

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

其中 $I^v(x)$ 和 $I^s(x)$ 在 HSV 颜色空间被表示为 $I^v(x) = \max_{c \in \{r, g, b\}} I^c(x)$ 和

$I^s(x) = (\max_{c \in \{r, g, b\}} I^c(x) - \min_{c \in \{r, g, b\}} I^c(x)) / \max_{c \in \{r, g, b\}} I^c(x)$ 。颜色衰减特征和场景深度成比例，可以用它容易地估计出透射率。

4. 色调差异：人们用原始图像 $I(x)$ 和它的半反图像

$I_{si}(x) = \max[I^c(x), 1 - I^c(x)]$ ($c \in \{r, g, b\}$) 之间的色调差异来检测雾。对于无雾图像，半反图像的三个通道的像素值不会全部反转，这就导致了原图像与半反图像之间的巨大的色调变化。在【29】中，色调差异特征被定义为：

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

其中角标 h 表示 HSV 颜色空间中的图像的色彩通道。根据等式(8)，透射率被反向传播到 $H(x)$ 。

3、 DehazeNet

第二章 A 中的大气散射模型表明透射率图的估计是复原无雾图像最关键的步骤。为了达到这个目的，我们提出了 DehazeNet，一个可训练的端到端系统，它可以学习原始雾图和对应的透射率图之间的特征关系。在本章中，我们阐述了去雾网络的设计，并且解释了这些设计是怎样和现有的去雾方法联系起来的。第五章将展示最后通过估计透射率图来复原图像的去雾操作。

A. DehazeNet 层的设计

DehazeNet 由级联的卷积和池化以及在接在它们之后的非线性激活函数组成。图 2 展示了 DehazeNet 的结构。DehazeNet 的各层和非线性激活函数被设计用来执行估计透射率的四个顺序操作，分别称为特征提取、多尺度映射、局部极值和非线性回归。下面我们详细描述这些设计。

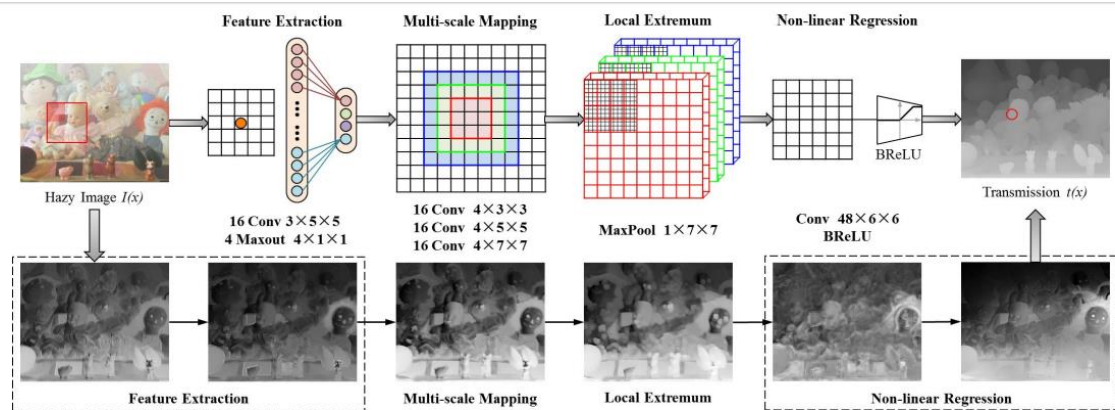


图 2. DehazeNet 的结构。DehazeNet 概念上包括四个连续的操作（特征提取、多尺度映射局部极值和非线性回归），结构上由 3 个卷积层一个最大池化，一个最大输出单元和 BReLU 激活函数组成。

1. 特征提取：为了克服去雾问题本身的病态性，现有的方法提出不同的假设，基于这些假设，人们能够提取出与雾有关的特征（如暗通道、色调差异、颜色衰减）。我们注意到提取这些与雾有关的特征和用一个合适的滤波器对输入的有雾图像进行卷积并进行非线性映射是相等的。受到那些与雾有关的颜色通道的极值处理的启发，我们选择了一个叫做最大输出单元的非线性映射来降低维度。最大输出单元是一个用于多层感知器或 CNNs 的简单非线性激活函数。当用于 CNNs 时，它通过对 k 个特征图采取像素级的最大化操作生成一个新的特征图。基于最大输出单元，我们设计的 DehazeNet 的第一层如下所示：

$$F_1^i(x) = \max_{j \in [1, k]} f_1^{i,j}(x), \quad f_1^{i,j} = W_1^{i,j} * I + B_1^{i,j}, \quad (9)$$

其中 $W_1 = \{W_1^{i,j}\}_{(i,j)=(1,1)}^{(n_1, k)}$ ，分别代表滤波器和偏置， $*$ 代表卷积操作。 $W_1^{i,j} \in R^{3 \times f_1 \times f_1}$ 是 $k \times n_1$ 个滤波器中的一个，其中 3 是输入图像的通道数量， f_1 是滤波器的空间尺寸（如表 1 所示）。最大输出单元映射了每一个 kn_1 维的向量到 n_1 维的，并且通过自动学习而不是现有方法中的启发式方法来提取与雾有关的特征。

2. 多尺度映射：【17】已经证明多尺度特征对去雾是非常有效的，它是在不同尺度下集中提取输入图片的特征。多尺度特征提取对于实现尺度不变性也是很有有效的。例如，GoogLeNet 的初始结构就采用了不同滤波器尺寸的平行卷积层，更好地解决了输入图像中的物体定位问题，也因此在此 ILSVRC14 【32】中获得了最好的实验结果。受到多尺度特征提取的启发，我们在 DehazeNet 的第二层使用了平行卷积操作，卷积滤波器的尺寸分别是 3×3 、 5×5 、 7×7 ，对于每一个尺度，都使用同样数量的滤波器。通常，第二层的输出被表示为：

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

其中 $W_2 = \{W_2^{p,q}\}_{(p,q)=(1,1)}^{(3, n_2/3)}$ 和 $B_2 = \{B_2^{p,q}\}_{(p,q)=(1,1)}^{(3, n_2/3)}$ 包含了 n_2 对参数，并被分成了三个组。 n_2 是第二层的输出维度， $i \in [1, n_2]$ 是输出特征图的索引。 $\lceil \cdot \rceil$ 表示向下取整， \setminus 表示取余操作。

3. 局部极值：为了实现空间不变性，视觉皮层上的复杂细胞会将来自简单细胞的刺激进行线性整合。【33】提出，一系列的池化操作可以描述复杂细胞的空间一致性的属性。根据 CNNs 【34】的典型结构，克服局部敏感性的方法就是对每个像素进行领域最大值操作。另外，局部极值与透射率的局部一致性这一假设相符，它普遍被用来消除透射率估计的噪声。因此我们在 DehazeNet 的第三层使用局部极值：

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

其中 $\Omega(x)$ 是以 x 为中心的 $f_3 \times f_3$ 的邻域，第三层的输出维度 $n_3 = n_2$ 。与 CNNs 中用来降低特征图的分辨率的最大池化操作不同，我们的局部极值操作是针对特征图的每一个像素，能够保持分辨率不变，取得更好的复原效果。

4. 非线性回归：在深度网络里，标准的非线性激活函数包括 Sigmoid 【35】和线性修正单元 ReLU。前者易出现梯度消失的情况，这会导致收敛慢或在训练中使局部优化不理想。为了克服梯度消失的问题，人们提出了 ReLU，它提供了离散化的特征值。然而，ReLU 主要针对分类问题，对于图像复原等回归问题并不是那么理想。尤其是 ReLU 只在输入小于 0 时才会限制它的值。这会导致响应溢出，尤其是在最后一层，因为对于图像复原而言，它的输出值必须在最大和最小值之间的一小段范围。为了达到这个目的，我们提出了一种双边修正线性单元 BReLU 作为激活函数来突破这种限制，如图 32 所示。受到 Sigmoid 和 ReLU 的启发，BReLU 保持了双边约束和局部线性的特性。基

于我们提出的 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 $W_4=\{W_4\}$ 包含了一个尺寸为 $n_3*f_4*f_4$ 的滤波器， $B_4=\{B_4\}$ 包含了一个偏置， $t_{\min}=0$ 和 $t_{\max}=1$ 是 BReLU 的界限值。根据(12)，激活函数的梯度为：

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

把上述四层串联起来组成 CNN，它基于可训练的端到端系统，所有与卷积层有关的滤波器和偏置都是要学习的神经网络的参数。我们注意到，这些层的设计都能与现有的去雾方法联系起来，在接下来的部分我们会详细介绍这些联系。

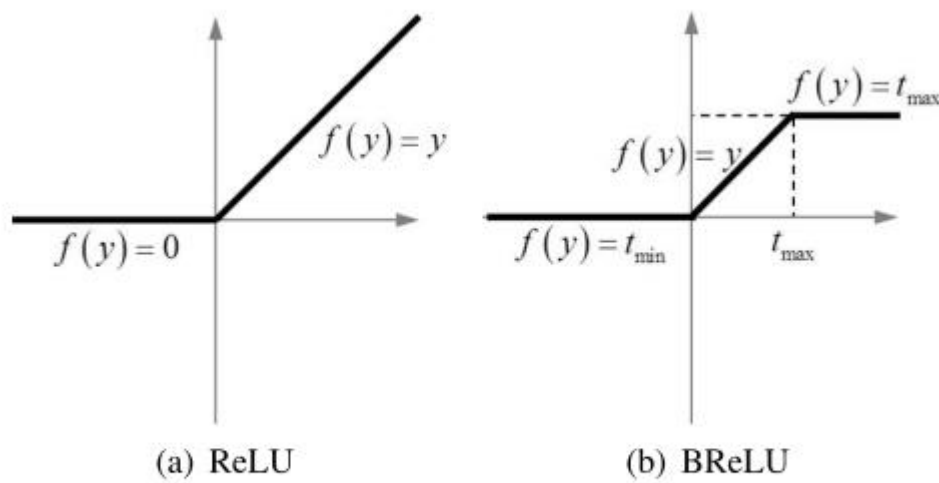


Fig. 3. Rectified Linear Unit (ReLU) and Bilateral Rectified Linear Unit (BReLU)

图 3. 修正线性单元（ReLU）和双边修正线性单元（BReLU）

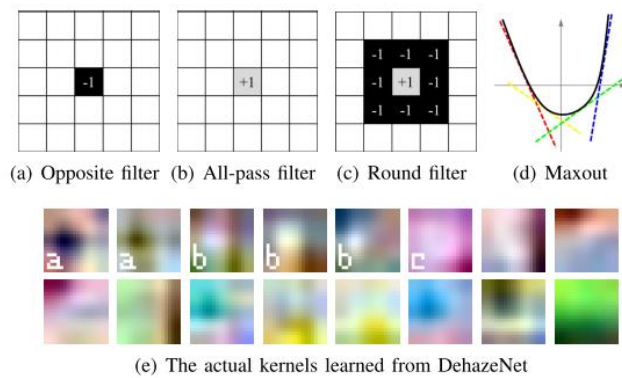


Fig. 4. Filter weight and Maxout unit in the first layer operation F_1

图 4. 第一层操作 F_1 中的滤波器权重和最大输出单元

B. 与传统去雾方法的联系

去雾网络第一层的特征 $F1$ 是用来提取与雾有关的特征的。拿暗通道【9】为例。如果权重 $F1$ 是一个反向滤波器（通道中心的值为-1 的稀疏矩阵，如图 4(a)所示）， $B1$ 是一个单位偏置，那么特征图的最大输出就等于颜色通道的最小值，这与暗通道【9】是相似的。同样的，当权重是环形滤波器， $F1$ 与最大对比度相似；当 $W1$ 同时包含全通滤波器和反向滤波器， $F1$ 等价于最大和最小特征图，它完成了从 RGB 到 HSV 空间的转换，等价于颜色衰减和色调差异特征。总而言之，根据图 4(e)展示的网络学习结果，几乎所有的与雾有关的特征都能被 $DehazeNet$ 的第一层提取到。另一方面，最大输出激活函数被认为是一种分段线性逼近任意凸函数。在本篇文章中，我们选择四个特征图中的最大值来近似一个随机的凸函数，如图 4(d)所示。

在一幅图像中，白色物体与高亮度低饱和度的浓雾场景相似。因此，几乎所有的雾估计方式都倾向于把白色物体当作远距离物体，这导致了透射率估计不准确。基于场景深度局部是一致的假设，局部极值滤波器通常被用来解决这个问题【9】，【18】，【7】。在 $DehazeNet$ 中，第三层的局部最大值滤波操作除去了局部估计误差。因此当透射率接近于 0 时，直接衰减项 $J(x)t(x)$ 也会接近于 0。此外直接复原场景辐射 $J(x)$ 极易受噪声影响。在 $DehazeNet$ 中，我们提出了 $BReLU$ 来限制透射率的值在 0 和 1 之间，因此消除了误差。值得注意的是， $BReLU$ 等价于传统去雾方法【9】、【18】中的边界约束。

C. 训练 $DehazeNet$

1. 训练数据：收集庞大的带有标签的数据来用于训练深度模型是很费时的【19】。对于训练 $DehazeNet$ ，这是更加困难的，因为有雾和相应的无雾场景图像通常是得不到的。我们基于物理的雾模型进行合成有雾图像【17】。

更加特别的，我们基于以下两个假设合成用于训练的有雾和无雾图像块：第一，图像内容与透射率无关；第二，透射率是局部一致的。这两个假设表明我们可以给一个图片块设置任意的透射率。给定一个没有雾的图片块 $J^p(x)$ ，大气光照 a ，和任意的 0 到 1 之间的透射率 $t(x)$ ，有雾图像就可以用方程 $I^p(x)=J^p(x)t+a(1-t)$ 来合成。为了减少不确定性，我们将大气光照设为 1。

在这项工作中，我们从网上收集无雾图片，然后随机地裁剪出 $16*16$ 的图片块。与【17】不同的是，这些图片不仅仅来自人们的日常生活，还有自然和城市景观，因为我们认为 $DehazeNet$ 的滤波器可以从这些各种各样的训练样本中进行学习。图 5 展示了我们收集到的一些无雾图片的例子。



Fig. 5. Example haze-free training images collected from the Internet

图 5. 从网上收集到的无雾的训练图片样本

2. 训练模型：在 DehazeNet 中，有监督的学习需要 RGB 和透射率之间的映射关系 F 。DehazeNet 通过最小化训练图片与真实透射率的损失来进行网络参数的学习。给定一系列的有雾图片块和他们的真实透射率，这些有雾图片块都是按照上面的方法用无雾图片合成的，我们使用均方差作为损失函数：

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

我们使用随机梯度下降法来训练 DehazeNet。我们使用 caffe 来建立我们的模型。表 1 总结了 DehazeNet 的详细配置和参数设置，其中包含了三个卷积层、一个最大池化层、第一层使用了最大输出单元，最后一个卷积层使用 BReLU 作为激活函数。

表 1. DehazeNet 模型参数

TABLE I. THE ARCHITECTURES OF THE DEHAZENET MODEL

Formulation	Type	Input Size	Num n	Filter $f \times f$	Pad
Feature Extraction	Conv	$3 \times 16 \times 16$	16	5×5	0
	Maxout	$16 \times 12 \times 12$	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

为了阐述 DehazeNet 的结构，我们分析了它的收敛性，并与其他先进的方法进行了比较。对于训练数据，我们从网上搜集到了一些无雾图片，并随机采样裁剪成了 10,000 张小的无雾图片块。对于每一个图片块，我们都使用位于 0 和 1 之间的 10 个透射率进行映射来产生 10 张有雾图片小块。这样一来，我们就合成列 10,000 张用于网络训练的图片块。在 DehazeNet 中，每一层的滤波器参数都使用高斯分布来随机初始化，所有的偏置被设为 0。学习率从 0.005 以每 100,000 次迭代下降一半的方式降到 3.125×10^{-4} 。基于以上的参数，DehazeNet 采用英伟达的 GTX780 GPU 进行训练。

根据 DehazeNet 估计出来的透射率和大气散射模型，无雾图像可以像传统方法一样被复原。因为第三层的局部极值的作用，透射率图中出现了块效应，为进一步优化透射率图，我们使用引导图像滤波【15】来平滑图像。根据方程(4)，在透射率图中，将透射率最小的 0.1% 个值作为边界值 t_0 ，我们将对 t_0 应位置的最大像素值作为大气光照的估计。给定透射率和大气光照，无雾图像可以很轻易地被恢复。为了方便，公式(1)被重新写成：

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

虽然 DehazeNet 是基于 CNNs，但它的轻量化结果可以有效地保证实时性能，并且可以运行在没有 GPU 的平台上。整个去雾框架在只有 CPU 的 Matlab2014a 上进行实验，处理一张 640×480 的图片只要大约 1.5 秒。

A. 模型和性能

在 DehazeNet 中，有两个重要的层是针对透射率进行特殊设计的，特征提取 F1 和非线性回归 F4。为了证明 DehazeNet 的高效性，我们将 DehazeNet 与两个传统的 CNNs (SRCNN[41]) 和 CNN-

L[23]进行了比较，这两个基准模型都拥有同样的三层网络结构。他们的参数个数分别是 8240、18400 和 67552。

1) 特征提取 F_1 中的最大输出单元： F_1 中激活单元是非线性降维的，用来近似传统的与雾有关的特征的提取。在图像处理过程中，对于发现原始的特性和减少噪声低维映射是一个核心步骤。例如，PCA【42】和 LDA【43】作为经典的低维线性映射被广泛使用于计算机视觉和数据最小化。在【23】中，一个非线性稀疏低维的映射和 ReLU 一起被用于高分辨率的重建任务中。作为一个特殊的低维映射，最大输出单元通过最大化特征图来发现雾图的先验知识。因此，下面的实验是被设计用来证实最大输出单元的有效性。根据【23】，线性单元映射了一个 16 维的向量到 4 维的向量，这等价于使用一个 $16*1*1$ 大小的滤波器。

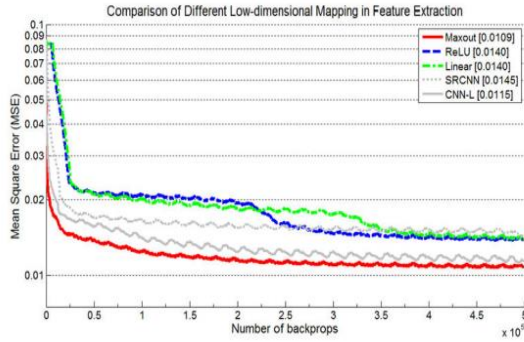


Fig. 6. The training process with different low-dimensional mapping in F_1

图 6. F_1 中不同维度映射的训练过程

图 6 展示了带有最大输出单元的 DehazeNet 的训练过程，并与 ReLU 和线性单元进行了比较。我们观察到，带有最大输出单元的网络的收敛速度比 ReLU 和线性单元要快。另外，括号内的值显示了收敛结果，相比于 ReLU 和线性单元，最大输出单元的性能提升了大约 0.3%。原因在于最大输出单元提供了几乎所有的与雾有关的特征的等价的函数，并且削弱了像 ReLU 一样的分段函数的劣势。

2) 非线性回归中的 BReLU: BReLU 是一个新的激活函数，它对于图像复原和重建是非常有效的。受到 ReLU 和 Sigmoid 的启发，BReLU 被设计成带有双边约束和局部线性的特性。双边约束使用了先验约束来减小搜索空间；局部线性克服了梯度消失的问题，并且获得了更高的精度。在对比实验中，使用 ReLU 和 Sigmoid 来代替非线性回归层的 BReLU。对于 ReLU， F_4 被写成 $F_4 = \max(0, W_4 * F_3 + B_4)$ ；对于 Sigmoid， $F_4 = 1 / (1 + \exp(-W_4 * F_3 - B_4))$ 。

图 7 展示了不同激活函数的训练过程。BReLU 比 ReLU 和 Sigmoid 拥有更快的收敛速率，尤其是在第一个 50,000 此迭代中。收敛精度显示，相比于 ReLU，BReLU 的性能提升了大约 0.05%；相比于 Sigmoid，提升了大约 0.2%。图 8 描绘出了预测的透射率与真实的透射率在不同的激活函数之间的关系。显然在 BReLU 的结果中预测的透射率在 45 度线附近。然而预测的透射率总是比实际的透射率要高，并且有些透射率超过了最大值 $t_{\max}=1$ 。由于函数的弯曲，预测的透射率与真实的透射率相差较大，接近于 0 和 1。BReLU 在测试集上的均方差是 0.0192，ReLU 和 Sigmoid 分别是 0.0128 和 0.0146。

B. 滤波器数量和尺寸

为了获得性能与参数尺寸之间最好的平衡，我们修改了去雾网络的参数。基于网络的缺省设置，

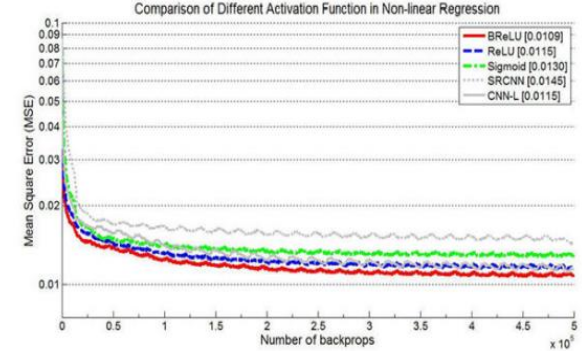


Fig. 7. The training process with different activation function in F_4

图 7. F_4 中不同激活函数下的训练过程

我们进行了两个实验：一个是使用非常多的滤波器，另一个是使用不同尺寸的滤波器。与第二节相似的，这些模型，都是在同样的数据集上进行训练。表 2 展示了不同参数设置下的训练或者测试的均方差。

表 2. 使用不同滤波器数量或尺寸的结果

TABLE II. THE RESULTS OF USING DIFFERENT FILTER NUMBER OR SIZE IN DEHAZENET ($\times 10^{-2}$)

Filter	Architecture	Train MSE	Test MSE	#Param
Number ($n_1 \cdot n_2$)	4-(16 \times 3)	1.090	1.190	8,240
	8-(32 \times 3)	0.972	1.138	27,104
	16-(64 \times 3)	0.902	1.112	96,704
F_2 Size ($f_1 \cdot f_2 \cdot f_3 \cdot f_4$)	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 ($f_1 \cdot f_2 \cdot f_3 \cdot f_4$)	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

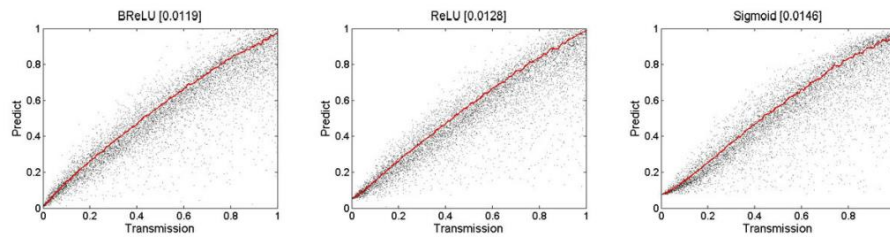


Fig. 8. The plots between predicted and truth transmission on different activation function in the non-linear regression F_4

图 8. 非线性回归 F_4 中不同激活函数下的预测透射率和真实值的曲线图

通常增大网络的宽度，能够获得更好的性能。显然增加滤波器的数量可以获得优越的性能。然而如果需要更快的去雾的速度，我们应该选择更小的网络，同时它也能比其他的方法获得更好的性能。在这篇文章中，下面的实验就采用了这个轻量化的网络。

另外我们测试网络在不同滤波器尺寸下的灵敏度。如表一所示的默认网络设置是时，我们首先分析了第二层中不同滤波器尺寸的影响。表 2 表明，更大尺寸的滤波器，可以获得更加丰富的结构信息，同时也能获得更好的结果。网络的第二层中采用了滤波器尺寸为 3、5、7 的多尺度特征映射，相对于单个尺度的滤波器来说他们实现了同样的测试均方差。更进一步，我们在第五节 D 中证明了，多尺度映射，能够提升尺度的鲁棒性。

然后我们又对网络在第三层和第四层下使用不同的滤波器尺寸进行了实验。在保持网络同样的接受域的同时调节第三层和第四层中的滤波器的尺寸。结果表明第四层非线性回归中使用更大尺寸的滤波器，可以增强拟合性，但是可能会导致过拟合。第三层的局部极值，可以提高测试集上的鲁棒性。因此我们发现了第三层和第四层的最好的滤波器尺寸是 5、6、7。

C. 合成图片块的定量结果

近些年去雾领域有三种基于学习框架的方法。在【18】中，去雾参数通过线性模型来学习，利用颜色衰减先验来估计场景深度。反向传播神经网络被用来发掘训练样本中的颜色和深度之间的内部关系。在【17】中，随机森林被用来获取与物有关的特征，用于无雾图像的复原。以上所有的方法和去雾网络都是用同样的方式来训练。利用无雾图片中随机采样得到的 2000 张图片块和十个随机透射率来产生 2 万张有雾图片用于测试。在同样的测试集上运行去雾网络来测量预测透射率和真实透射率的均方差。【9】是一种典型的去雾方式，通常被用来作为比较的基准。

表三展示的测试机上的预测透射率和真实透射率的均方差。去雾网络实现的最好的得分，为 0.0119。将图片的特征值分类，这样一来就打断了图像内容和与雾有关的特征的联系。去网络涉及到

的，图像内容信息，对于天空和白色物体的透射率的估计是非常有用的。在随后的实验中得到了令人满意的结果。但是在这个实验中，由于线性回归模型中一些超出范围的事大于 1 或者小于 0，模型的性能较差。

表 3. 合成图片块的预测透射率和真实透射率的均方差

TABLE III. MSE BETWEEN PREDICTED TRANSMISSION AND GROUND TRUTH ON SYNTHETIC PATCHES

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



Fig. 9. Synthetic images based on Middlebury Stereo Datasets and DehazeNet results

图 9. 基于 Middlebury 立体数据集的合成图片和 dehazeNet 结果

D.合成图片的定量结果

为了验证在完整图片上的效果，我们在合成的有雾图片上测试了 DehazeNet，并且和其他方法进行了比较，这些图片都来自于已知深度的立体图片。图 9 中，有雾图像是利用等式(1)通过无雾的立体图像合成的，然后通过 DehazeNet 还原成无雾图像。

为了定量地评估这些方式，我们针对每一对无雾图像和去雾结果图片的差异使用了一系列的评

估标准。除了广泛使用的 MSE 和 SSIM 标准，我们还使用了另外的评估方法，分别是 WPSNR 和 PSNR。我们定义 OPE 为传统的方法，运行标准参数，报告性能的平均测量结果。在表 4 中，DehazeNet 通过 OPE 和所有有雾图片与其他 6 中先进的方法进行了比较。令人兴奋的是，DehazeNet 虽然是用均方差进行优化，但他比其他方法都表现出了更好的性能。

表 4. 合成图片下 MSE、SSIM、PSNR 和 WSNR 的平均结果

TABLE IV. THE AVERAGE RESULTS OF MSE, SSIM, PSNR AND WSNR ON THE SYNTHETIC IMAGES ($\beta = 1$ AND $\alpha = 1$)

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

表 5. 不同散射系数、图片尺度和大气光照下合成图片的均方差

TABLE V. THE MSE ON THE SYNTHETIC IMAGES BY DIFFERENT SCATTERING COEFFICIENT, IMAGE SCALE AND ATMOSPHERIC AIRLIGHT

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

去雾效果对于雾的浓度是很敏感的，不同散射系数下的性能它的性能也是相差很大的。因此，我们提出了一种关于散射系数下的去雾鲁棒性的方法，称为散射鲁棒性评估。如表 5 所示，CAP [18] 在散射系数为 0.75 时表现出了很好的性能，但当雾的浓度增加时去雾效果逐渐降低。原因在于，CAP 是基于预测场景深度和假设散射系数为 1 来估计透射率的。在 [17] 中，使用 200 棵树来建立随机森林用于非线性回归，展示出了很好的系数鲁棒性。然而，随机森林里每一个像素的大的计算量限制了它的实际应用的能力。至于 DehazeNet，透射率是通过 F1 中的非线性激活函数估计得到的，所以对于散射系数表现出了极好的鲁棒性。

由于雾和光源的颜色补偿，大气光并不是纯白的。我们提出了一种大气光鲁棒性评估方法来分析不同大气光照下的去雾结果。虽然 DehazeNet 是在大气光照设为 1 的情况下进行训练的，对于其他不同的大气光照值，它仍然表现出了极好的鲁棒性。特别的，当阳光，雾的比例为[1.0,1.0,0.9]时，DehazeNet 比其他方法表现得更好。因此，DehazeNet 也能用来去除白色模糊，如图 10 所示，这种白色模糊是围绕光源的白色圆环。



Fig. 10. Image enhancement for anti-halation by DehazeNet

图 10. DehazeNet 的抗白晕的图像增强

在真实世界的应用中通常会出现可视视野转换和图像缩放的情况。尺度鲁棒性评估，用来分析

不同尺度的影响。与在 OPE 中表现同样好的方法比较，我们从 0.4 到 1.0 之间选择了 4 个尺度系数来生成不同尺寸的图像用于尺度鲁棒性评估。在表五中，DehazeNet 由于使用了 F2 中的多尺度映射，在不同的尺度下都表现出了极好的鲁棒性。在 CAP【18】、DCP【9】和 ATM【39】中只使用了单个尺度导致了不同尺度上的预测精度差异。当图像收缩时，过度的大尺度邻域处理将会丢失图像的细节。因此，DehazeNet 中的多尺度映射提供了不同的滤波器来融合多尺度特征，并且它在所有的尺度下都获得了最高的得分。

在许多情况下，传感器或摄像头电路都会随机产生噪声，进而导致估计误差。我们也讨论了不同程度的图像噪声对我们的方法的影响。作为一个基本的噪声模型，带有标准偏差的附加高斯白噪声(AWG)被用来评估噪声鲁棒性(NRE)。得益于 F1 中最大输出抑制和 F3 中的局部极值，DehazeNet 比其他方法表现出了更好的噪声鲁棒性。RF【17】在许多方面的评估都表现很好，唯独噪声评估方面表现较差，主要由于图片块的特征值通过分类打断了透射率和图像内容的联系，这也会放大异常值的影响。

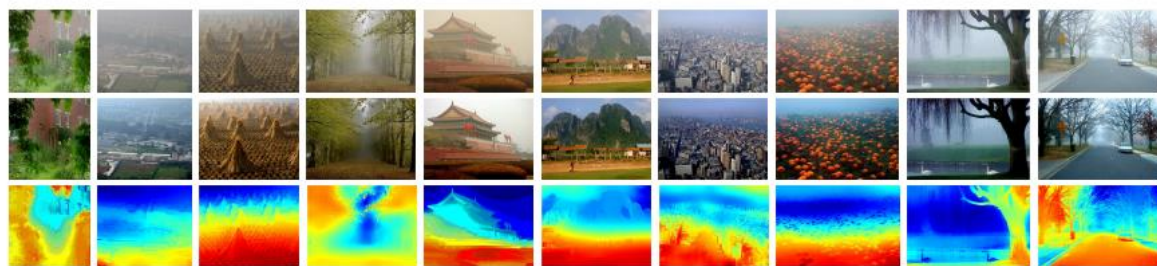


Fig. 11. The haze-free images and depth maps restored by DehazeNet

图 11. 有雾图像和被 DehazeNet 恢复的深度图

E. 真实世界图片上的定性结果

图 11 显示了被 DehazeNet 复原的去雾结果和深度图，更多的结果和比较可以在这里找到：<http://caibolun.github.io/DehazeNet/>。因为所有的去雾算法在一般的室外图片上都能获得较好的结果，所以很难用可视化的方式进行排序。为了比较他们，本文聚焦于在【9】、【17】、【18】相关研究中五张被大家所认同的有挑战性的图片。这些图片都有大的白色或灰色区域，很难处理，因为大部分现有的去雾算法对于白色很敏感。图 12 显示了六种先进的去雾算法的定性比较结果。图 12 展示了原图和各种算法的去雾结果。图 12(h)展示了 DehazeNet 的去雾结果。

有雾图片的天空区域去雾是很难的，因为云和雾都是有着同样的大气散射模型的相似的自然现象。如前面三张图片所示，在(b-d)的结果中大部分的雾都被去掉了，并且场景和物体的细节被很好的复原了。然而，去雾结果都受到了天空区域的过度增强的严重影响。总的来说，这些图片的天空区域都是比真实的图片更暗或者是过饱和以及失真的。雾通常只存在于大气表层，因此天空区域基本不需要处理。基于学习框架，CAP 和 RF 避免了天空的颜色失真，但是由于无内容的回归模型导致了非天空区域的增强较差。DehazeNet 善于找到天空区域和保持颜色不变，并且在其他区域保证了较好的去雾影响。原因在于图片块的特征可以被 DehazeNet 的隐藏层学习到，然后影响到天空区域的去雾。

因为基于先验的透射率估计是一种类型的数据，对于特定的图片可能失效。第四和第五张图片注定是【9】中的失败案例。当场景物体和大气光相似时，基于先验估计到的透射率是不真实的。因为暗通道和这些物体有着相近的明亮值，FVR 和 BCCR 都是基于 DCP 的，存在一个固有的过度估计透射率的问题。从回归模型中进行学习的 CAP 和 RF 不会受到过饱和的影响，但是会在较大距离时低估雾的浓度。与这六种方法相比，DehazeNet 的结果避免了图像的过饱和，并且由于它的非线性

回归得到了较好的去雾效果。

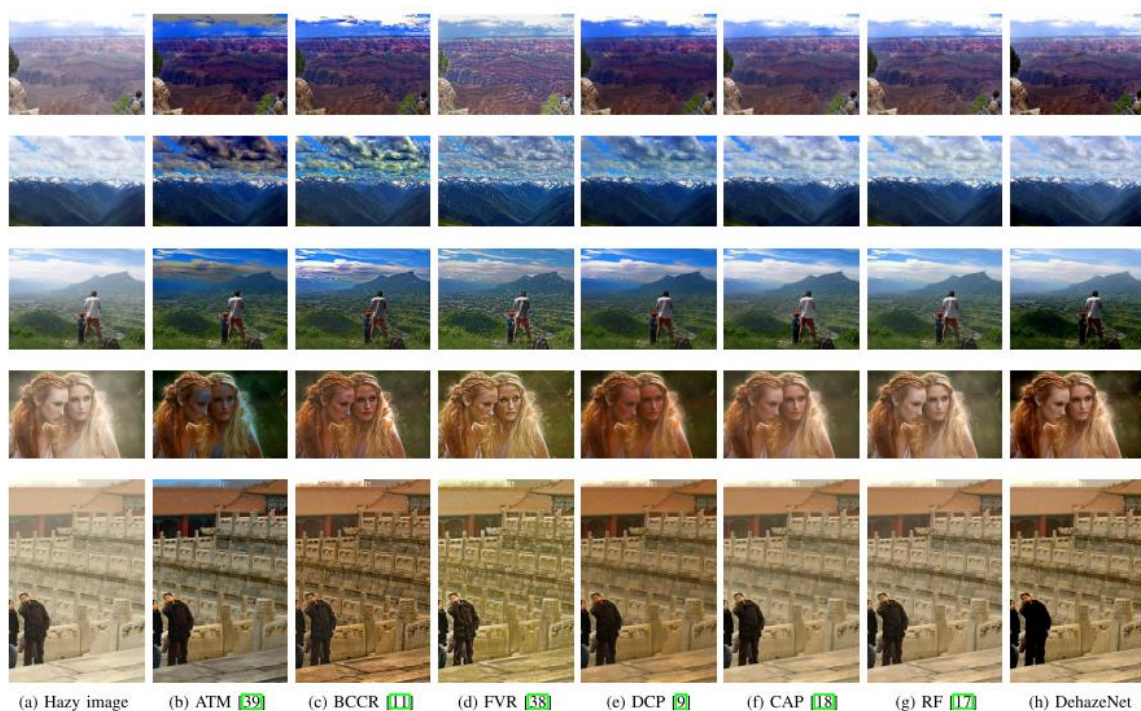


Fig. 12. Qualitative comparison of different methods on real-world images.

图 12. 不同方法的真实世界图片结果的定性比较

5、结论

在这篇文章中,我们为图像去雾提出了一种新型的深度学习算法。受传统的与雾相关的特征和去雾方法的启发,我们指出透射率的估计过程可以用一个经过特殊设计的可训练的端到端系统重新表示出来,其中特征提取层和非线性回归层均在经典的 CNNs 结构上做了改进。在第一层 F1 中,我们证明了最大输出单元与以前的方法具有类似的效果,同时在提取雾相关特征方面效率更高。在最后一层 F4 中,一种新型的激活函数 BReLU 替代了 ReLU 和 Sigmoid 以保持图像复原的双边限制和局部线性。得益于轻巧的网络结构,DehazeNet 同时实现了高效率和显著的去雾效果,比目前的所有算法都要优秀。

但是,需要指出的是,尽管我们成功地将 CNN 应用于图像去雾,在我们的工作中,仍然有许多需要改进的地方。例如,大气光不能被视作全局不变的,而应该与透射率一起被网络学习。而且,我们认为,大气散射模型也可以在更深层的网络中被学习得到,这样的话,雾图和清晰图像之间的端到端映射就可以被直接优化得到,而不用经过透射率估计这一步骤。在今后的研究中,我们会着重解决这些问题。