

PD-Net: 基于金字塔池化模块改进的Dehaze-Net

宋翼洋, 刘 刚

成都信息工程大学应用数学学院, 四川 成都

收稿日期: 2021年9月18日; 录用日期: 2021年10月11日; 发布日期: 2021年10月19日

摘 要

对单幅雾天图像进行去雾是一个困难的任务, 本文提出了一种名叫PD-Net的神经网络模型用于单幅图像去雾。该模型在Dehaze-Net模型中引入了金字塔池化模块且额外添加残差块。金字塔池化模块增强了模型对全局信息的提取, 残差块有效地抑制了梯度消失现象。在仿真实验中, 本文使用RESIDE数据集中的室内合成有雾图片作为实验数据集, 定量分析比较了各种去雾方法对实验数据集的去雾结果, PD-Net模型展现了良好的性能。在真实实验中, 对RESIDE数据集中的室外真实有雾图片进行去雾分析。真实实验的结果表明, 相较于其他算法, PD-Net模型在大面积的天空区域及图片的细节处有更好的效果。

关键词

去雾, Dehaze-Net, 金字塔池化模块, PD-Net

PD-Net: Improved Dehaze-Net Based on Pyramid Pooling Module

Yiyang Song, Gang Liu

Department of Applied Mathematics, Chengdu University of Information Technology, Chengdu Sichuan

Received: Sep. 18th, 2021; accepted: Oct. 11th, 2021; published: Oct. 19th, 2021

Abstract

It is a difficult task to remove haze from a single input image. In this paper, we present a neural network model called PD-Net for dehazing single image. In this model, Pyramid Pooling Module is introduced into Dehaze-Net model, and Residual blocks are added. The pyramid pooling module enhances the extraction of global information, and the residual block effectively suppresses the disappearance of gradient. In the simulation experiment, we use the indoor synthetic foggy pictures in the RESIDE set as the experimental set, quantitatively analyze and compare the defogging results

of various defogging methods on the experimental set, and the PD-Net model shows good performance. In the real experiment, we defog and analyze the images which is the outdoor real foggy pictures in the RESIDE set. The results of real experiments show that compared with other algorithms, PD-Net model has better effect in large-area sky area and details of picture.

Keywords

Defogging, Dehaze-Net, Pyramid Pooling Module, PD-Net

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

图像中蕴藏了大量的信息,在大数据时代的背景下,图像是人类不可或缺的传递信息的媒介。伴随着工业的发展,大气污染愈发的严重,导致了雾霾天气的频繁出现。在雾霾天气下,由于光的反射作用以及空气中雾霾颗粒的存在,使光在传入图像采集设备时能量衰减,导致采集的图像色彩暗淡、模糊不清、对比度低,大量的信息因为雾霾被掩盖。利用图像去雾技术消除雾霾,可以在一定程度上恢复图像的信息。图像去雾作为计算机视觉领域的低层次视觉任务,对后续的语义分割、目标识别等视觉任务具有重要影响。

目前国内外的去雾方法研究方向主要分为三类:基于图像增强的方法、基于物理模型的方法以及基于深度学习的方法。第一类基于图像增强的方法能够很好地改善图片的对比度信息、恢复图片的色彩信息,但是普遍对细节的处理较差,例如 Pizer 等在文献[1]中提出的自适应直方图均衡化的方法,Adelmann 在文献[2]中提出的图像同态滤波的方法,Finlayson 在文献[3]中提出的单尺度 Retinex 图像增强算法。第二类基于物理模型的方法,这里的物理模型主要指 Narasimhan 等在文献[4]中提出的大气散射模型,这类算法的关键在于依据一定的先验知识估计大气散射模型中的两个未知参数透射率和大气光照值,例如凯明等在文献[5]中提出的暗通道先验去雾算法。第三类基于深度学习的方法近年来得到了越来越多的重视,该类方法主要通过搭建深度学习模型来实现图像去雾。蔡博仑等在文献[6]中提出了首个用于图像去雾的卷积神经网络模型 Dehaze-Net。Dehaze-Net 将先验信息融入网络模型的特征提取层中,通过卷积层和激励层提取图像信息,输出图像的透射率图,再根据大气散射模型恢复无雾图像。此外,基于深度学习的图像去雾方法近几年也得到了发展,Ren 等在文献[7]中设计了一种多尺度卷积神经网络模型 MSCNN 用于图像去雾, Li 等在文献[8]中提出了首个端到端的卷积神经网络模型 AOD-Net。

Dehaze-Net 模型在图像去雾中表现了良好的性能,经该模型去雾后的对比度高,无明显噪点,较为符合人眼的感知;且 Dehaze-Net 模型结构小巧,运行仅需占用很少的内存,方便移植到小型的设备。但在复杂的场景中,Dehaze-Net 模型对边界信息的提取并不充分,去雾后图像的细节略有瑕疵;且在大面积的天空区域,一些去雾后的图片出现了颜色失真现象。本文针对这些问题,改进了 Dehaze-Net,提出基于金字塔池化模块[9]改进的 Dehaze-Net (Pyramid pooling Dehaze-Net, PD-Net)模型。PD-Net 模型添加了金字塔池化模块(Pyramid Pooling Module, PPM)并额外引入了残差块。添加的 PPM 增强了模型聚合图像的全局信息的能力;引入的残差块抑制了梯度消失现象,使模型更快的达到收敛。实验结果表明,PD-Net 模型去雾后的图像在大面积的天空区域更能保留原本的色彩,整体更符合人眼的感知。

2. 模型背景

2.1. 大气散射模型

在图像去雾领域, 雾天图像的成像原因一般通过大气散射模型[4]来解释, 大气散射模型具体地表示为:

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

其中, $I(x, y)$ 为在雾天采集到的图像在位置 (x, y) 处的像素值, $J(x, y)$ 为未经过空气中雾霾粒子散射的真实图像在位置 (x, y) 处的像素值, A 为大气光照值, $t(x, y)$ 为场景透射率, 可以表示为:

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

其中 β 是大气散射系数, $d(x, y)$ 为场景的深度。

将大气散射模型进行变换, 可以得到逆大气散射模型:

$$J(x, y) = \frac{I(x, y) - A(1 - t(x, y))}{t(x, y)}. \quad (3)$$

由逆大气散射模型(3)可知, 图像去雾的关键就是求得两个未知参数大气光照值 A 以及场景透射率 $t(x, y)$ 的值, 带入公式即可恢复无雾清晰图像 $J(x, y)$ 。对于逆大气散射模型中三个未知数一个方程的情况, 一般情况下只需求 $t(x, y)$ 的值, 而 A 值可以先验理论进行估计, 再对无雾清晰图像 $J(x, y)$ 进行求解。

2.2. Dehaze-Net 模型

随着深度学习的发展, 基于传统去雾算法中的先验特征, Cai 等人[6]提出了 Dehaze-Net 模型。该模型用于估计场景透射率 $t(x, y)$, 再通过暗通道先验理论[5]估计大气光照值 A , 将估计的 A 和 $t(x, y)$ 带入逆大气散射模型即可恢复无雾清晰图像。

Dehaze-Net 由多层卷积层和池化层组成, 在其中一些层后采用了适当的非线性激活函数。Dehaze-Net 采用特征提取, 多尺度映射, 局部极值和非线性回归估计场景透射率 $t(x, y)$, 具体结构如图 1。

Dehaze-Net 的第一层特征提取层采用了 Maxout 激活函数, Maxout 激活函数从通道方向进行最大值滤波, 通过提取最大值形成新的特征; 受 GoogLeNet [10]中 inception 架构的启发, Dehaze-Net 的第二层多尺度映射层使用不同大小的卷积核进行并行卷积运算, 用于提取信息; 根据卷积神经网络(Convolutional Neural Networks, CNN) [11]的经典架构, 使用最大池化可以增强模型鲁棒性, 此外, 最大池化与透射率局部一致的假设性一致, 故 Dehaze-Net 的第三层采用最大池化运算; 由于模型的输出为估计的透射率 $t(x, y)$, 应该控制在 $(0, 1)$, 为此, 提出了 BReLU 激活函数(图 2)作为 Dehaze-Net 的最后一层。

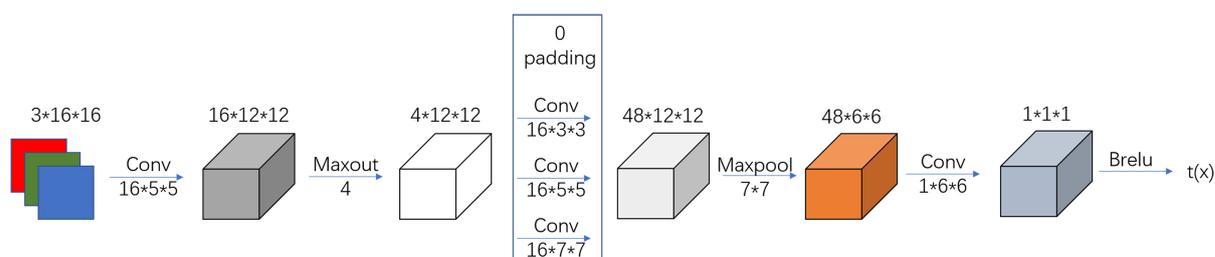


Figure 1. The architecture of Dehaze-Net

图 1. Dehaze-Net 结构图

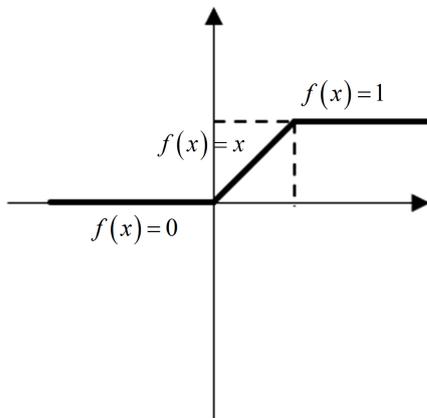


Figure 2. Bilateral rectified linear unit
图 2. 双向修正线性单元图

Dehaze-Net 模型的训练数据使用无雾图像合成的有雾图像, 模型使用学习的合成有雾图像与透射率之间的映射关系, 对真实有雾图像的透射率进行估计, 再根据暗通道先验理论以及逆大气散射模型恢复无雾清晰图像, Dehaze-Net 模型去雾的一般流程见图 3。

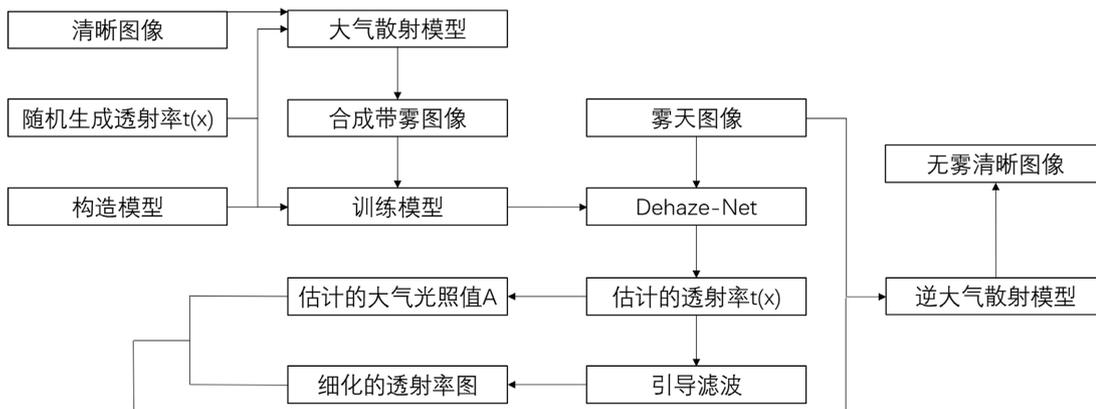


Figure 3. Process of using Dehaze-Net to remove fog
图 3. Dehaze-Net 去雾流程图

3. 本文模型

本文以 Dehaze-Net 为基础进行改进, 提出 PD-Net 模型。PD-Net 在 Dehaze-Net 的第一层特征提取层和第三层局部极值池化层之间通过跳跃连接加入残差块结构; 将 Dehaze-Net 的第二层多尺度映射层的张量填充替换为反射填充; 在第四层激活层之前加入金字塔池化模块。本章第一节介绍金字塔池化模块, 第二节介绍残差块, 第三节介绍 PD-Net 模型。

3.1. 金字塔池化模块

PPM [9]可以进一步减少不同子区域间空间信息的丢失, 它多尺度的池化操作, 能够聚合不同尺度、不同子区域间的信息。在 Zhao 等[9]提出的金字塔场景解析网络(Pyramid Scene Parsing Network, PSPNet)中, PPM 就表现出了良好性能。金字塔池化生成的不同层次的特征图最终被拼接起来, 送入后续的网络以进行进一步的信息提取。金字塔池化模块的具体结构见图 4。

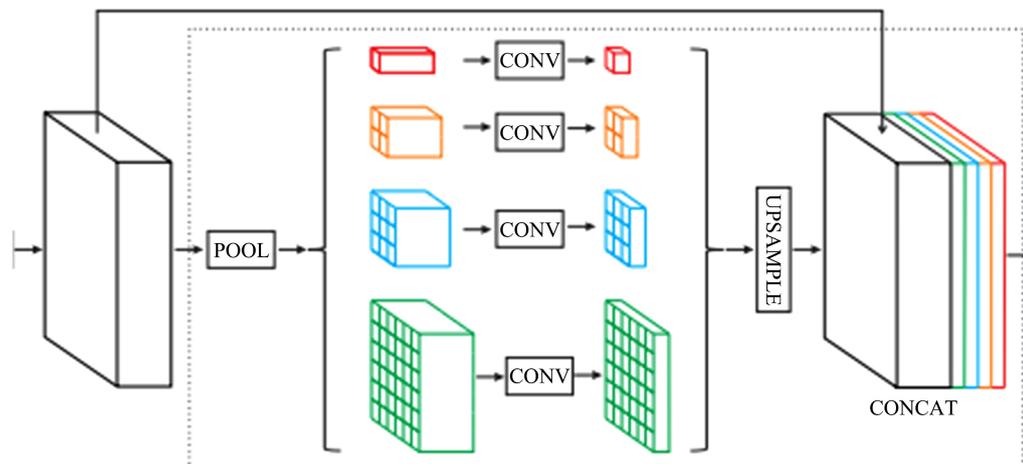


Figure 4. The architecture of PPM
图 4. 金字塔池化模块图

3.2. 残差块

在残差块提出之前, 梯度消失一直是深度学习里一个难以解决的问题。随着网络层数的增加, 一些复杂的模型出现了梯度消失现象, 导致模型的拟合效果反而不如简单的模型, 严重时甚至无法训练模型。因此, 神经网络的层数一直较少。直到何凯明提出了 ResNet [12], 神经网络的层数达到了 152 层。ResNet [12] 基于 VGG19 [13] 网络进行了修改, 并通过跳跃连接加入了残差块结构, 残差块结构如图 5。残差块的跳跃连接抑制了在深层网络中最容易出现的梯度消失现象; 当网络出现退化现象时, 跳跃连接能够使网络更容易学习到恒等映射。

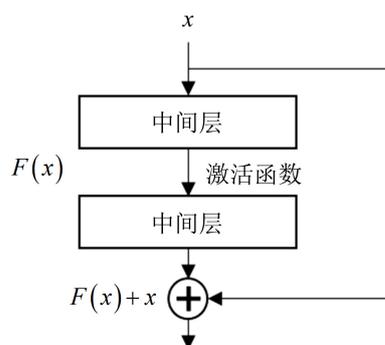


Figure 5. Residual block
图 5. 残差块结构图

3.3. PD-Net

针对 Dehaze-Net 在复杂场景中对全局信息的提取不充分问题, 本文提出了 PD-Net。PD-Net 在 Dehaze-Net 的基础上加入了 PPM [9] 并额外添加了残差块 [12] 结构。PPM 聚合了图像不同区域间信息, 使模型能够更充分的提取图像全局信息。残差块结构有效的抑制了梯度消失和网络退化现象, 减小了模型的训练难度。在 Dehaze-Net 的第二层多尺度映射层中使用反射填充替换 0 值填充; 加强了边界信息的提取。

具体的结构见图 6, 各层的输入及输出见表 1。

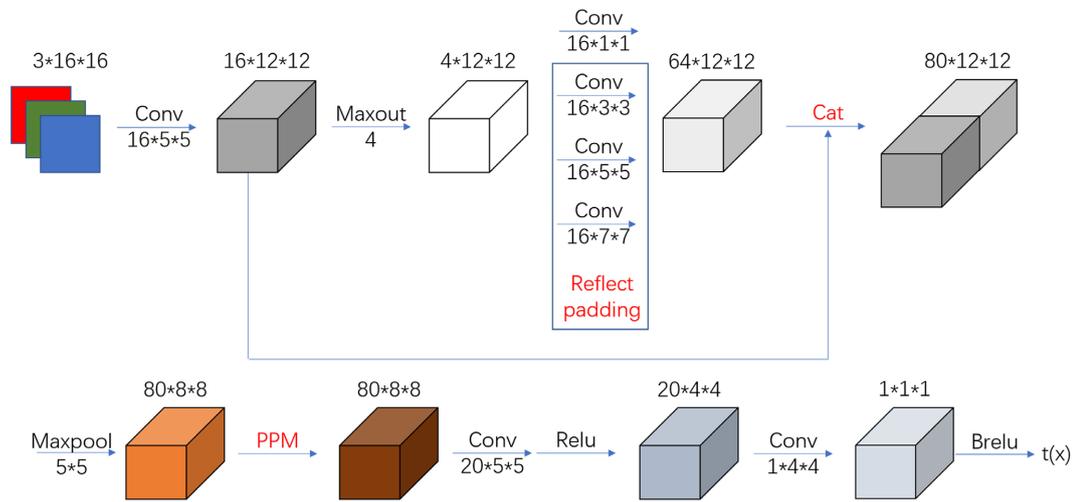


Figure 6. The architecture of PD-Net
图 6. PD-Net 结构图

Table 1. The parameter settings of the PD-Net
表 1. PD-Net 各层结构表

Type	Input size	num	Kernel size	pad	Output size
Conv1_1	3*16*16	16	5*5	0	16*12*12
Maxout	16*12*12	4	/	0	4*12*12
Conv2_1	4*12*12	16	1*1	0	16*12*12
Conv2_2	4*12*12	16	3*3	1	16*12*12
Conv2_3	4*12*12	16	5*5	2	16*12*12
Conv2_4	4*12*12	16	7*7	3	16*12*12
Maxpool	80*12*12	/	5*5	0	80*8*8
PPM	80*8*8	/	/	/	80*8*8
Conv3_1	80*8*8	20	5*5	0	20*4*4
ReLu	20*4*4	/	/	/	20*4*4
Conv4_1	20*4*4	1	4*4	0	1*1*1
BReLU	1*1*1	/	/	/	1*1*1

PD-Net 模型使用合成有雾图片进行训练。训练数据集包括加雾后的室内及室外的清晰无雾图片，使用大气散射模型对图片加雾，对每张图片分别加十次雾，将(0, 1)平均分为十个区间，每次加雾的 $t(x)$ 从十个区间里随机取值。模型的训练步骤如下：

第一步：制作数据集、构造 PD-Net 模型、构建损失函数和优化器。

第二步：将数据集中的合成带雾图片输入模型。

第三步：初始化模型参数，计算第二步中得到的输出与真实的 $t(x)$ 之间的损失函数。

第四步：根据损失函数，使用反向传播算法计算模型各参数的梯度，通过优化器更新模型的参数。返回第二步，直至损失函数最小。

Table 2. Quantitative results on RESIDE in terms of full-reference image quality assessment
表 2. 全参考图像质量评估指标在 RESIDE 数据集上的定量结果

评估指标	CLAHE [14]	SSR [3]	DCP [5]	Dehaze-Net [6]	PD-Net
SSIM	0.6101	0.6217	0.8447	0.8003	0.8187
PSNR	13.12	8.54	17.46	16.99	17.44



Figure 7. Defogging results on indoor synthetic dataset
图 7. 室内合成带雾图片上的去雾结果图

4. 实验结果和分析

本文所有的算法及模型是在 GPU 型号为 MX450、显存为 2GB, 系统为 Windows10 x64 的笔记本电脑上, 通过 Anaconda 平台的 Python 语言, 使用 PyTorch、OpenCV、Numpy 等第三方库实现的。

在本章中将通过定量的分析来验证 PD-Net 模型在图像去雾上的性能。本章使用的评估指标包括全参考图像质量评估指标 PSNR 和 SSIM [15], PSNR 可以表示为:

$$PSNR = 10 \times \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right), \quad (4)$$

其中 MSE 指原图像与处理后图像之间的均方误差。

此外, 本文还将 PD-Net 与其他去雾方法应用到 RESIDE 数据集, 并对去雾结果进行比较, 分析各个算法的优劣性, 包括基于对比度受限的自适应直方图均衡化去雾算法(Contrast Limited Adaptive Histogram Equalization, CLAHE) [15], 基于单尺度 retinex 去雾算法(Single Scale Retinex, SSR) [3], 基于暗通道先验的去雾算法(Dark Channel Prior, DCP) [5]以及基于深度学习的 Dehaze-Net 去雾模型。

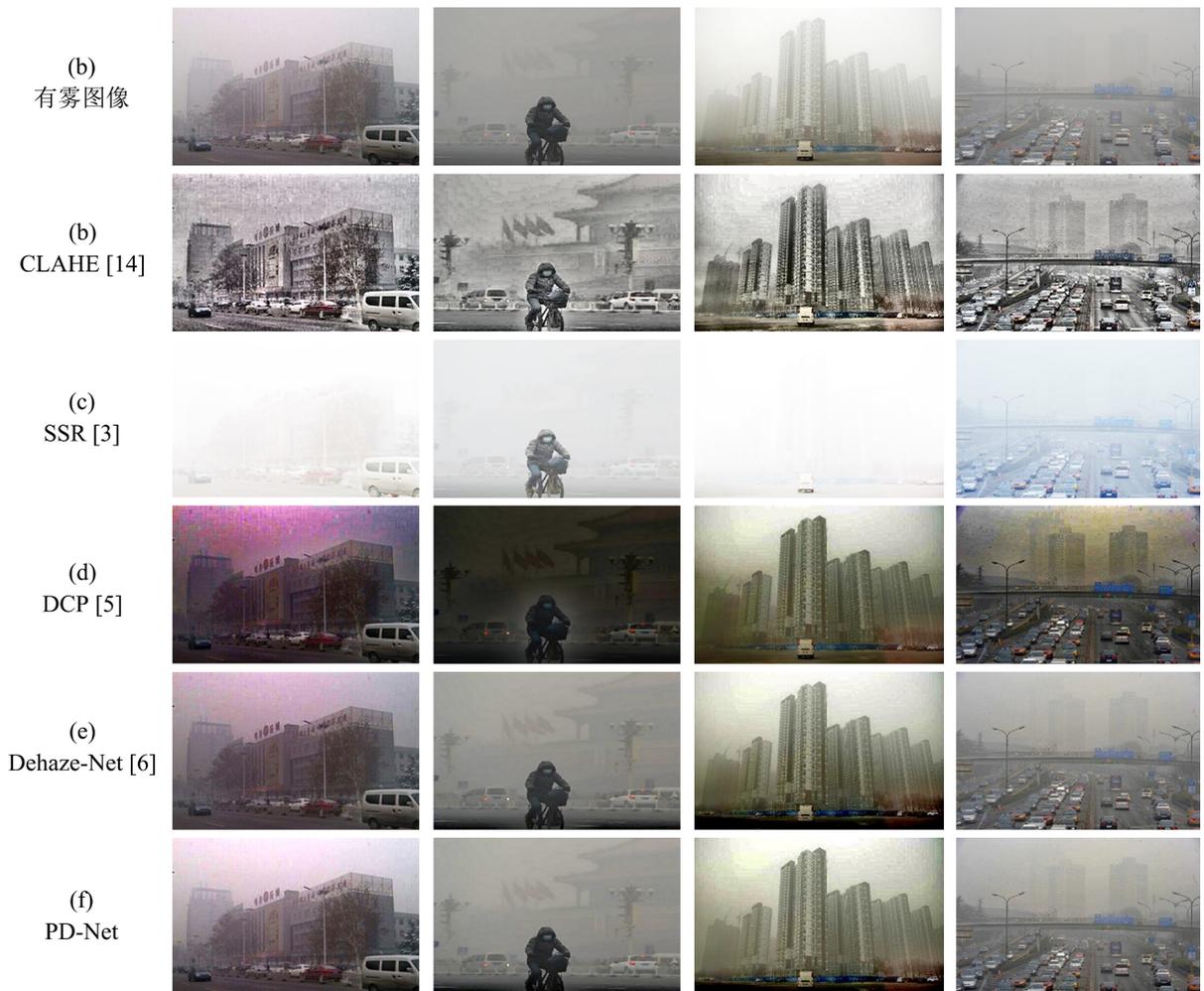


Figure 8. Defogging results on the outdoor real fog picture
图 8. 室外真实有雾图片上的去雾结果图

表 2 展示了去雾方法的平均 PSNR 值和 SSIM 值, 图 7 展示了各方法以及 PD-Net 模型在 RESIDE 数据集中的室内合成带雾图片上的去雾效果。从表 2 可以看到, PD-Net 模型在实验数据集上得到了良好的平均 SSIM 值和 PSNR 值, 仅次于 DCP 算法, 但从实际效果上看, PD-Net 模型的去雾效果明显优于 DCP 算法。从图 7(c) 可以看到, 经过 CLAHE 算法去雾后的图片中的墙体和桌面存在大量噪声, 图片不清晰; 从图 7(d) 可以看到, 经过 SSR 算法去雾后的图片整体偏白, 去雾并不彻底, 并且有过度曝光现象; 从图 7(e) 可以看到, 经过 DCP 算法去雾后的图片对比度低, 在图中的墙体处有颜色失真现象, 并在图中的柜门处存在噪声; 从图 7(f) 可以看到, Dehaze-Net 模型去雾后的图片细节恢复不错, 但也存在颜色失真现象。通过放大观察图 7(g) 的细节, 可以看到经 PD-Net 模型去雾后的图片细节更加完美, 并且图片清晰明亮, 没有明显的颜色失真现象, 更加符合人眼的感知。

为了进一步分析 PD-Net 模型的去雾效果, 本文还对 RESIDE 数据集中的室外真实有雾图片进行了去雾, 见图 8。可以看到, CLAHE 算法去雾后产生了许多噪声; SSR 算法导致了图片中出现了大量的空白区域; DCP 算法对大面积的天空区域的处理效果不好, 有颜色失真现象; Dehaze-Net 模型和 PD-Net 模型的效果较好, 但 Dehaze-Net 模型去雾后图片中的部分天空区域也出现了颜色失真现象。综上所述, PD-Net 模型去雾效果的优势是明显的。

5. 结语

本文以 Dehaze-Net 为基础, 引入金字塔池化模块及残差块, 构造了 PD-Net。实验证明, PD-Net 模型对天空区域和边缘细节的处理效果较好, 处理后的图片清晰明亮, 质量高。经对比实验证明, PD-Net 模型去雾效果较 Dehaze-Net 模型有明显的提升, 并优于其余算法。

致 谢

感谢胡建成老师在研究中提供的帮助, 也感谢匿名的审稿人提出的宝贵意见。

基金项目

本研究由国家自然科学基金项目 11901065 资助。

参考文献

- [1] izer, S.M., Amburn, E.P., Austin, J.D., et al. (1987) Adaptive Histogram Equalization and Its Variations. *Computer Vision Graphics & Image Processing*, **39**, 355-368. [https://doi.org/10.1016/S0734-189X\(87\)80186-X](https://doi.org/10.1016/S0734-189X(87)80186-X)
- [2] Adelman, H.G. (1998) Butterworth Equations for Homomorphic Filtering of Images. *Computers in Biology & Medicine*, **28**, 169-181. [https://doi.org/10.1016/S0010-4825\(98\)00004-3](https://doi.org/10.1016/S0010-4825(98)00004-3)
- [3] Finlayson, G.D., Hordley, S.D. and Drew, M.S. (2002) Removing Shadows from Images Using Retinex. *Proceedings Color Imaging Conference*, Scottsdale, 12-15 November 2002, 73-79. https://doi.org/10.1007/3-540-47979-1_55
- [4] Narasimhan, S.G. and Nayar, S.K. (2003) Contrast Restoration of Weather Degraded Images. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **25**, 713-724. <https://doi.org/10.1109/TPAMI.2003.1201821>
- [5] He, K., Jian, S., Tang, X.O., et al. (2011) Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **33**, 2341-2353. <https://doi.org/10.1109/TPAMI.2010.168>
- [6] Cai, B., Xu, X., Jia, K., et al. (2016) Dehaze-Net: An End-to-End System for Single Image Haze Removal. *IEEE Transactions on Image Processing*, **25**, 5187-5198. <https://doi.org/10.1109/TIP.2016.2598681>
- [7] Ren, W., Si, L., Hua, Z., et al. (2016) Single Image Dehazing via Multi-Scale Convolutional Neural Networks. In: *European Conference on Computer Vision*, Springer, Cham, 154-169. https://doi.org/10.1007/978-3-319-46475-6_10
- [8] Li, B., Peng, X., Wang, Z., et al. (2017) AOD-Net: All-in-One Dehazing Network. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 4780-4788. <https://doi.org/10.1109/ICCV.2017.511>
- [9] Zhao, H., Shi, J., Qi, X., et al. (2017) Pyramid Scene Parsing Network. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 6230-6239. <https://doi.org/10.1109/CVPR.2017.660>

- [10] Szegedy, C., Wei, L., Jia, Y., *et al.* (2015) Going Deeper with Convolutions. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 1-9. <https://doi.org/10.1109/CVPR.2015.7298594>
- [11] Lecun, Y. and Bottou, L. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <https://doi.org/10.1109/5.726791>
- [12] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [13] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition.
- [14] Zuiderveld, K. (1994) Contrast Limited Adaptive Histogram Equalization. In: *Graphics Gems IV*, Morgan Kaufmann, Burlington, 474-485. <https://doi.org/10.1016/B978-0-12-336156-1.50061-6>
- [15] Wang, Z., Bovik, A.C., Sheikh, H.R., *et al.* (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, **13**, 600-612. <https://doi.org/10.1109/TIP.2003.819861>