

用于单张图像去雨滴的轻量级网络^①



蔡林纹, 王冠

(天津大学 数学学院, 天津 300350)

通讯作者: 蔡林纹, E-mail: cailinwen@tju.edu.cn

摘要: 图像去雨是图像低等级任务中的热点问题, 去雨滴又是图像去雨中很重要的一种情况, 附着在玻璃或相机镜头上的雨滴会显著降低场景的可见性. 因此, 去除雨滴将有助于许多计算机视觉应用, 特别是户外监控系统和智能驾驶系统. 本文提出了一种用于单张图像去雨滴的轻量级网络算法 (PRSEDNet), 该网络算法采用递归计算, 运用卷积长短期记忆网络 (Convolutional LSTM network) 和特征提取模块来提取特征, 通过与原图像结合来去除雨滴, 最终获得高质量的无雨滴清晰图. 实验结果表明, 我们的 PRSEDNet 与现有的基于深度学习的去雨滴算法相比, 在能达到高效的去雨滴性能的同时, 有更少的参数量且计算效率高.

关键词: 雨滴去除; 深度学习; 神经网络; 递归计算

引用格式: 蔡林纹, 王冠. 用于单张图像去雨滴的轻量级网络. 计算机系统应用, 2021, 30(8): 201-206. <http://www.c-s-a.org.cn/1003-3254/8032.html>

Lightweight Network for Single Image Raindrop Removal

CAI Lin-Wen, WANG Guan

(School of Mathematics, Tianjin University, Tianjin 300350, China)

Abstract: Image de-raining is a hot issue in the low-level tasks of images, in which raindrop removal is critical. Raindrops adhering to glass or camera lenses will significantly reduce the visibility of the scenes. Therefore, removing raindrops will benefit various computer vision tasks, especially outdoor surveillance systems and intelligent driving systems. In this study, we propose a lightweight network (PRSEDNet) to remove raindrops from a single image. With recursive computation, this network uses a convolutional long short-term memory network and a feature extraction module to extract features. In combination with the original image, the final high-quality clear image without raindrops is obtained. The experimental results show that PRSEDNet, compared with the existing deep learning-based raindrop removal algorithms, possesses few parameters and high computational efficiency and can achieve efficient raindrop removal.

Key words: raindrop removal; deep learning; neural network; recursive computation

下雨是一种常见的天气状况, 雨不仅会影响人类的视觉感知, 而且会影响室外的计算机视觉系统. 在雨天, 雨线 (rain streak) 和附着在相机镜头上的雨滴 (raindrop) 都会显著降低拍摄图像的质量. 因此, 雨线或雨滴去除问题引起了广泛的研究. 近些年来, 研究者在去除雨线上已经取得了很好的成果. 然而, 这些方法不能直接

用于去除雨滴, 因为雨滴的形状和物理成像过程与雨线完全不同. 图像上的雨线一般是呈条纹状均匀分布在整张图像上, 且雨线的方向是一致的. 与雨线相比, 雨滴呈现出更加不规则的几何形状, 在图像上的分布没有规律, 使得雨滴去除问题更具有挑战性.

一些研究者们基于雨滴的物理成像过程或几何形

^① 收稿时间: 2020-11-10; 修改时间: 2020-12-12, 2020-12-25; 采用时间: 2021-01-06; csa 在线出版时间: 2021-07-31

态来模拟雨滴的分布. Kurihata 等^[1] 试图通过主成分分析 (PCA) 来检测雨滴, 但当场景中的物体与雨滴相似时, PCA 无法有效地提取雨滴的特征, 而且容易获得错误的检测结果. Yamashita 等^[2] 先是通过测量玻璃和立体相机之间的距离来检测雨滴, 用其它图像区域替换该雨滴位置来达到去除雨滴的目的, 但这种方法是在假设其它图像的相同位置没有雨滴覆盖的前提下进行的, 实际中这一假设有较大的局限性. 之后, Yamashita 等^[3] 又通过移动相机拍摄出的图像序列来替代立体相机拍摄出的图像.

近些年来, 随着深度学习的流行和广泛应用, 有一些研究者们开始尝试利用神经网络来解决雨滴去除问题. Eigen 等^[4] 建立了一个三层网络来提取合成图像中雨滴的特征. 虽然这种方法在雨滴小而稀疏的情况下效果较好, 但它不能去除大而密集的雨滴. Qian 等^[5] 提出了一种可以同时检测和去除雨滴的方法 Attentive-GAN, 该方法先利用一个小型递归网络检测雨滴在图像上的位置并生成相应的注意力图 (attention maps), 然后这些注意力图进一步用于后续网络处理, 以提高去雨滴的性能. 虽然该方法可以很好地去除单张图像中的雨滴, 然而其网络结构包含的参数相对较多, 这限制了它在计算资源有限的实际应用中的潜在价值. 因此, 设计有效且高效的雨滴去除算法对于广泛的实际应用是很有必要的. 由于 PReNet^[6] 在去除雨线方面是一种很好的轻量级网络基本模型 (Baseline), 但直接将其应用于雨滴去除任务, 得到的去雨滴结果并不好.

本文基于 PReNet 提出了一种用于单张图像雨滴去除的网络 PRSEDNet, 与现有的基于深度学习的去雨滴方法相比, 可以获得相近或更优的去雨滴效果. 特别地, PRSEDNet 显著降低了网络的参数量. 实验结果验证了我们的 PRSEDNet 可以在达到高效的去雨滴性能

的同时, 有更少的参数量, 能获得和 Attentive GAN 相近的去雨效果, 并且比轻量级网络 PReNet 表现得更好.

1 本文方法

为了更好地细化分析本文方法, 在本节中, 我们首先介绍了雨滴去除问题的基本理论模型, 其次对我们使用的网络结构进行了描述, 最后给出了我们使用的损失函数.

1.1 基本理论模型

Attentive GAN 中将雨滴图像建模为:

$$I = (1 - M) \odot B + R \tag{1}$$

其中, I 是带雨滴的图像, B 和 R 分别为背景图层和雨滴图层, M 是一个二值变量, 当 $M=1$ 时, 代表此像素属于雨滴区域, 当 $M=0$ 时, 代表该像素属于背景区域. 但其对 M 的真值采用了如下的设置: 当 $x \geq 30, M=1$; 当 $x < 30, M=0$, 其中 x 为图像像素 (取值范围为 0-255). 该方法简单地将阈值设置为 30, 这种方式在实际应用中有一定的局限性. 为了让模型有更好的泛化能力, 我们不使用雨滴位置的信息, 即上述的 M , 而是将雨滴去除问题看作是二个图层的分离问题, 建模如下:

$$I = B + R \tag{2}$$

其中, I, B, R 的定义同上.

1.2 网络结构

本文提出的去雨滴网络 PRSEDNet 如图 1 所示, 由两个卷积层, 卷积长短期记忆网络和特征提取模块组成. 即 1 个卷积层 f_{in} 用于接收输入图像, 经过卷积长短期记忆网络 f_{lstm} 和特征提取模块 f_{block} 来提取特征, 1 个卷积层 f_{out} 用于输出去除雨滴后的清晰图像. 因此, 第 t 阶段 PRSEDNet 可以用如下公式表述:

$$x^t = y + f_{out} \left(f_{block} \left(f_{lstm} \left(f_{in} \left(x^{t-1} \right) \right) \right) \right) \tag{3}$$

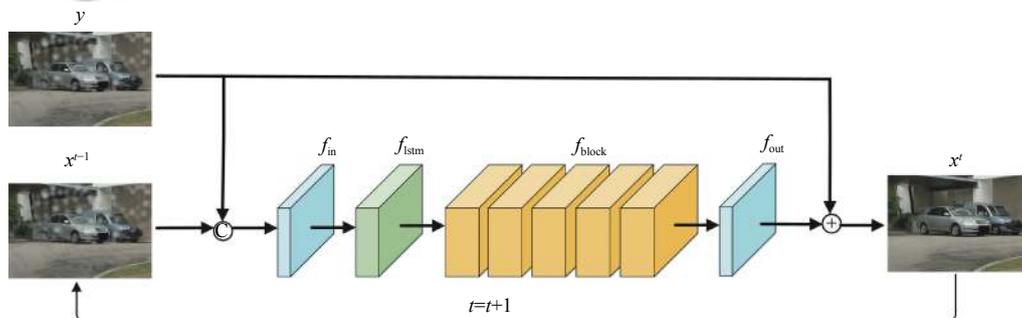


图 1 PRSEDNet 的网络结构

最大阶段数为 T , 除特殊说明, 网络中所有的卷积层的卷积核大小都设置为 3, 步幅 (stride) 和填充 (padding) 都设置为 1, 激活函数使用 ReLU^[7]. 在整个过程中, 都没有上下采样的操作, 卷积核的其它参数设置都是为了让特征图的分辨率不会发生改变, 旨在恢复出的清晰图不会有图像内容细节的损失. 下面, 我们将给出每一层的实现细节.

f_{in} : 该部分仅由一个卷积层组成. 3 通道的雨滴图 y 和 3 通道的输出图 x^{t-1} 按通道数连接在一起组成一个 6 通道的输入, 将 6 通道输入送给卷积核大小为 3 的卷积层, 输出一个 32 个通道的特征图. 其中, x^0 初始化为 y .

f_{stm} : 这部分参考 PReNet 的网络结构, 本文也使用 ConvLSTM^[8] 中所提出的 Convolutional LSTM 结构. Convolutional LSTM 接收一个 32 个通道的特征图, 最终输出一个通道数仍为 32 的特征图.

f_{block} : 此部分为特征提取模块, 一共由 5 个 RSED Block 组成, 每个 RSED Block 可以分为 3 个部分, 即两个卷积层和一个注意力机制模块, 具体结构如图 2 所示. 接收到的特征图先通过一个卷积核大小为 3 的空洞卷积, 该空洞卷积的填充 (padding) 和扩张 (dilation) 都设置为 10. 之后再经过一个卷积核大小为 3 的卷积层 (该卷积不是空洞卷积, 步幅和填充均设置为 1), 最后再通过一个注意力机制模块, 我们在实验中采用 SENet^[9] 中的 SE 模块作为注意力机制模块, SE 的具体结构如图 3 所示, H 和 W 分别代表特征图 X 的长和宽, C 表示通道数, r 是放缩因子. 实验中我们将 C 和 r 分别设置为 32 和 16. 每个 RSED Block 中的输入和输出都是 32 个通道, 在特征提取的整个过程中我们和文献 [10] 一样使用了残差学习^[11].

f_{out} : 该部分作为最终的输出层仅由一个卷积层组成. 该卷积层的卷积核大小为 3, 且没有激活函数. 它将前面特征提取模块输出的 32 个通道的特征作为输入, 输出一个 3 通道的图像, 并将其加到原始雨滴图 y 上, 生成最终的去雨滴图像.

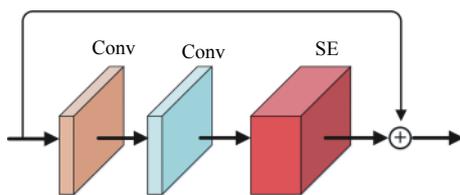


图 2 RSED block 的具体结构

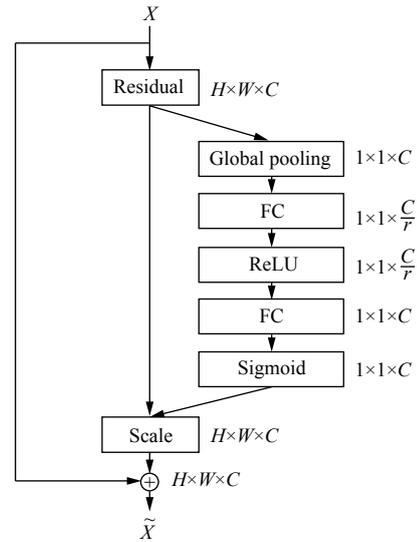


图 3 SE 模块的具体结构

1.3 损失函数

PRSEDNet 有 T 个阶段, 通过实验我们发现只需对网络最终的输出 x^T 进行监督. 我们不再简单地使用 SSIM 损失函数, 而是将 L_1 损失和 L_{W-SSIM} 损失^[12] 结合起来, 最终采用的损失函数如下:

$$Loss = \alpha L_1(x^T, gt) + \beta L_{W-SSIM}(x^T, gt) \quad (4)$$

其中, α 和 β 分别代表两种损失在整个损失中的权重系数, gt 是用于监督训练的无雨滴清晰图像, L_{W-SSIM} 是在 SSIM 损失函数中运用哈尔小波^[13], 使 SSIM 损失函数接收到的不再是整个图像, 而是经过哈尔小波变换计算出的各个图像切片. 通过这种方式, 我们可以在损失函数中改变图像分辨率, 以有效地抑制最终输出图像出现不希望的伪影.

2 实验分析

2.1 实验方法

实验采用全监督方法, 用雨滴数据集 (参考 2.2 节) 中雨滴图和相对应的无雨滴图 (ground truth) 形成训练对, 然后在我们提出的网络结构上进行端到端 (image-to-image) 的训练. 旨在学习到雨滴图到无雨滴图的映射或残差.

2.2 数据集和评价指标

实验采用 AttentiveGAN 中 Qian 等创建的雨滴数据集^[5] 来训练和测试我们的算法. 该数据集一共有 861 对雨滴图和清晰图用于训练, 含有两个测试集, test A 和 test B. 其中 test A 是 test B 的子集, test A 包含 58

对质量优良且对齐的图像, test B 包含 249 对图像, 但 test B 中有部分图像没有完全对齐, 用来计算评价指标会得到不准确的结果, 为了便于定量比较并精准计算评价指标, 我们只在 test A 上测试我们的结果, 并将结果与其他去雨滴算法进行比较。

评价指标我们采用最常用的衡量图像质量的两个值, 峰值信噪比 (PSNR)^[14] 和结构相似度 (SSIM)^[15]。PSNR 和 SSIM 值越大, 表示图像失真越小, 质量越好。

2.3 数据预处理和实验条件

在整个网络训练中, 我们设置了合适的卷积核以及相应的步幅和填充以使处理图像的分辨率不发生变化。为了加快网络的训练且减小显存占用, 我们将训练的 861 对图像全都裁剪为 100×100 分辨率的图像块, 然后将所有的图像块输入到网络进行图像端到端的训练。

我们设置网络的最大阶段数 T 为 6。损失函数 (4) 中的权重系数 α 和 β 分别设置为 0.15 和 0.85。实验采用 PyTorch 框架, 并在单张 NVIDIA GTX 1080Ti GPU 上进行训练。梯度下降的优化器采用 Adam^[16], Batchsize 固定为 18。整个实验训练 100 个 epoch, 初始学习率设为 0.001, 当达到第 30, 50 和 80 个 epoch 时, 学习率变为前一阶段的一半。

表 1 几种算法在测试集 test A 上的 PSNR 与 SSIM 平均值比较

算法	Eigen ^[4]	PReNet ^[6]	AttentiveGAN ^[5]	PRSEDNet (本文)
PSNR	20.9486 (↑11.1155)	30.7259 (↑1.3382)	32.7349 (↓0.6708)	32.0641
SSIM	0.8476 (↑0.1198)	0.9578 (↑0.0096)	0.9692 (↓0.0018)	0.9674

表 2 几种算法的参数数量比较

算法	PReNet ^[6]	AttentiveGAN ^[5]	PRSEDNet (本文)
参数量	168 963	6 244 362	169 603

图 4 给出了本文和其他算法在测试集 test A 上的去雨滴效果图, 为了便于观察, 我们从每张图中选取了 3 个矩形区域进行放大并排列在相应图的下方。如图 4 中的第一幅图所示, Eigen 算法只能处理很小且稀疏的雨滴, 当遇到大而密集的雨滴时, 其不能达到去除雨滴的效果, PReNet 能去掉部分雨滴, 但其去除雨滴不彻底 (参考红色和绿色特写框)。从图 4 中第 2 幅图的 3 个特写框可以看出, 四种算法去除雨滴都不彻底, 但 Eigen, PReNet, AttentiveGAN 三种算法会在得到的清晰图中保留更明显的雨滴印迹, 相比之下我们的算法可以得到更优的效果。图 4 中第 3 幅图的绿色特写框

2.4 数值和视觉结果

为了量化去雨滴的性能, 表 1 给出了 Eigen, PReNet, AttentiveGAN, PRSEDNet 四种算法的结构相似度与峰值信噪比的数值比较 (其中上下箭头分别代表提升和不足, 数值是我们的结果和其它算法结果相减得到的)。其中 Eigen 的结果是用其作者训练好的模型直接在测试集 test A 上进行测试, 而 AttentiveGAN 的作者在 GitHub 上并未给出训练代码, 只给出了其训练好的模型。我们直接使用其提供的模型在测试集 test A 上进行测试。虽然 PReNet 的作者在 GitHub 上提供了训练好的模型, 但由于它是用于处理雨线去除问题, 其提供的模型不能直接用于雨滴去除问题, 我们便按照其作者的默认设置参数在训练集上重新训练了 PReNet 来进行测试。可以看到, 我们的结果远远优于 Eigen 的结果, 和我们的基准模型 PReNet 相比有很大的提升, 同时也能达到和 AttentiveGAN 相近的结果。我们还给出了 PReNet, AttentiveGAN, PRSEDNet 三种算法的参数量比较, 具体见表 2, 值得注意的是, 我们比 PReNet 只多了很少的参数, 但去雨滴效果却提升很多, 而且我们能得到和 AttentiveGAN 很相近的结果, 但我们的参数量只有其参数量的 2.7% 左右。

反映了 PReNet 在生成清晰图像时会产生不希望出现的伪影, 而我们的算法在损失函数中运用了哈尔小波变换, 能有效地抑制该问题的出现。通过视觉效果图的对比可以看出, 我们所提出算法的去雨滴效果远远优于 Eigen 算法的效果, 同时可以达到不输于 AttentiveGAN 算法的效果, 并且比同为轻量级网络的 PReNet 效果更好。

为了更好地证明我们算法的有效性, 图 5 给出了整个实验流程中 PRSEDNet 和 PReNet 随 epoch 变化的 PSNR 和 SSIM 值。从图 5 中可以看出, PRSEDNet 的 PSNR 和 SSIM 值在实验中均大于 PReNet 的 PSNR 和 SSIM 值, 这进一步说明了 PRSEDNet 在去雨滴任务中比 PReNet 更有效。另一方面, PRSEDNet 的 PSNR 和 SSIM 曲线最终趋于平缓, 这也说明了我们算法的在实验中是可以收敛的。

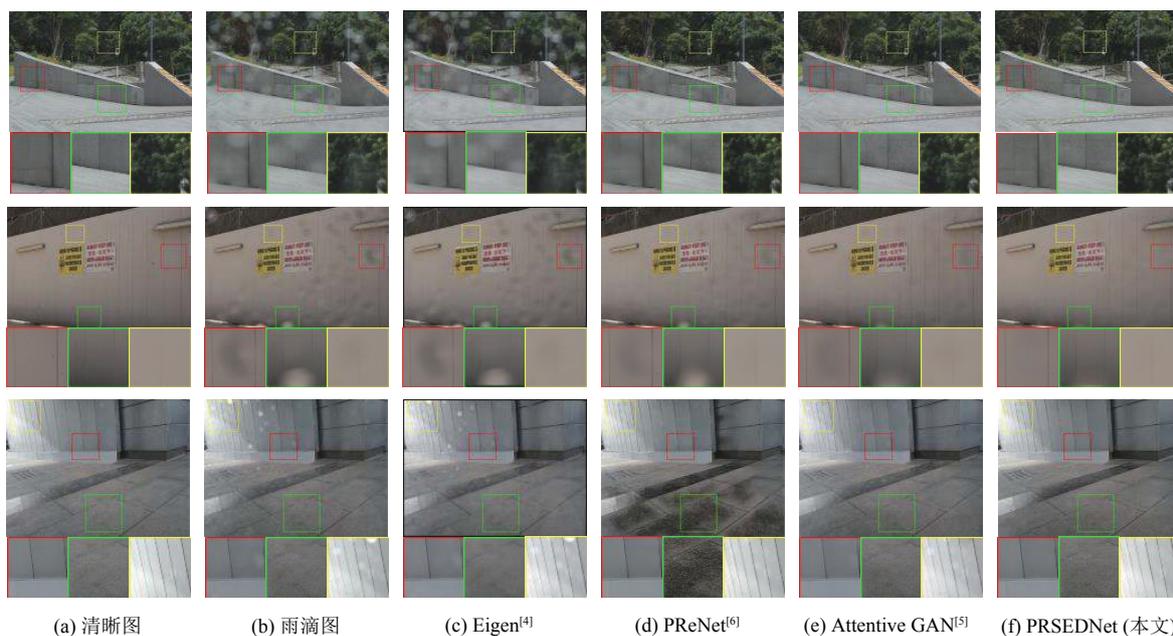


图4 不同算法在测试集 test A 上的去雨滴效果对比

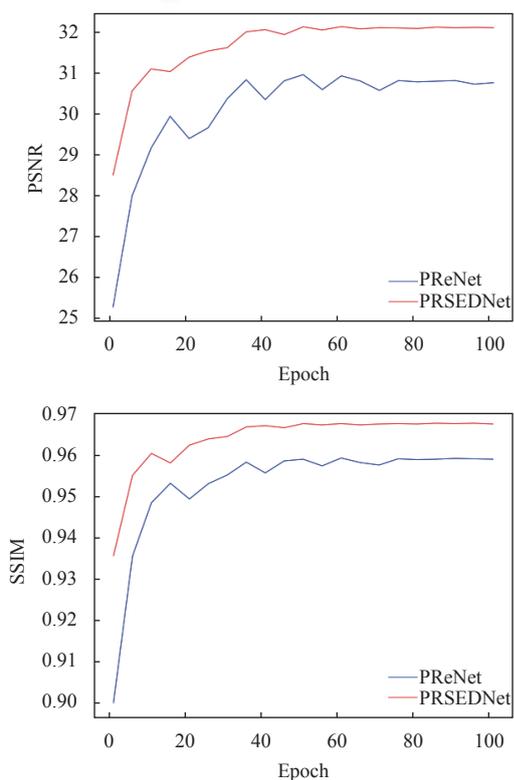


图5 训练过程中 PSNR 和 SSIM 值的变化

3 总结

本文提出了一种用于单张图像去雨滴的轻量级网

络算法. 实验结果表明, 与当前流行的去雨滴方法 Eigen 和 AttentiveGAN 相比, 我们的 PRSEDNet 可以达到更优或相近的去雨滴效果, 和我们的基本模型轻量级网络 PReNet 相比, 无论是在去雨滴清晰图的视觉感观上, 还是在 PSNR 和 SSIM 数值上, 我们的去雨滴效果都更好. 此外, 我们所提出的轻量级网络也可以应用于其他图像恢复任务中, 以平衡图像的恢复质量和相应的计算开销.

参考文献

- 1 Kurihata H, Takahashi T, Ide I, *et al.* Rainy weather recognition from in-vehicle camera images for driver assistance. IEEE Proceedings. Intelligent Vehicles Symposium. Las Vegas, NV, USA. 2005. 205–210.
- 2 Yamashita A, Tanaka Y, Kaneko T. Removal of adherent waterdrops from images acquired with stereo camera. IEEE/RSJ International Conference on Intelligent Robots and Systems. Edmonton, AB, Canada. 2005. 400–405.
- 3 Yamashita A, Fukuchi I, Kaneko T. Noises removal from image sequences acquired with moving camera by estimating camera motion from spatio-temporal information. IEEE/RSJ International Conference on Intelligent Robots and Systems. St. Louis, MO, USA. 2009. 3794–3801.
- 4 Eigen D, Krishnan D, Fergus R. Restoring an image taken through a window covered with dirt or rain. Proceedings of

- the IEEE International Conference on Computer Vision. Sydney, NSW, Australia. 2013. 633–640.
- 5 Qian R, Tan RT, Yang WH, *et al.* Attentive generative adversarial network for raindrop removal from a single image. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 2482–2491.
 - 6 Ren DW, Zuo WM, Hu QH, *et al.* Progressive image deraining networks: A better and simpler baseline. IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA. 2019. 3932–3941.
 - 7 Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel. 2010. 807–814.
 - 8 Shi XJ, Chen ZR, Wang H, *et al.* Convolutional LSTM network: A machine learning approach for precipitation nowcasting. Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, QC, Canada. 2015. 802–810.
 - 9 Hu J, Shen L, Albanie S, *et al.* Squeeze-and-Excitation networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011–2023. [doi: [10.1109/TPAMI.2019.2913372](https://doi.org/10.1109/TPAMI.2019.2913372)]
 - 10 Fu XY, Huang JB, Zeng DL, *et al.* Removing rain from single images via a deep detail network. IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 1715–1723.
 - 11 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770–778.
 - 12 Yang HH, Yang CHH, Tsai YCJ. Y-Net: Multi-scale feature aggregation network with wavelet structure similarity loss function for single image dehazing. IEEE International Conference on Acoustics, Speech and Signal Processing. Barcelona, Spain. 2020. 2628–2632.
 - 13 Stéphane M. A WaveletTour of Signal Processing: The Sparse Way. Academic Press, 2009.
 - 14 Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment. Electronics Letters, 2008, 44(13): 800–801. [doi: [10.1049/el:20080522](https://doi.org/10.1049/el:20080522)]
 - 15 Wang Z, Bovik AC, Sheikh HR, *et al.* Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 2004, 13(4): 600–612. [doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861)]
 - 16 Kingma D, Ba J. Adam: A method for stochastic optimization. 3rd International Conference for Learning Representations. San Diego, CA, USA. 2015.