

基于改进的残差网络的单幅图像去雨方法

崔 艳, 蔡瑞琪*, 于小亿

山西师范大学物理与信息工程学院, 山西 临汾

Email: cuiyan8013@163.com, *270027505@qq.com, yuriiac@hotmail.com

收稿日期: 2020年12月5日; 录用日期: 2020年12月28日; 发布日期: 2021年1月5日

摘 要

雨水是室外比较常见的天气现象。在雨水条件下, 光的散射和能见度下降, 导致计算机视觉算法会受到不同程度的干扰。为了进一步降低雨水给计算机视觉算法带来的影响, 本文提出一种基于改进的残差网络的单幅图像去雨方法。该方法结合了图像处理传统方法和深度学习的方法, 首先通过传统方法对有雨图像进行稀疏化, 将稀疏化后的图像作为输入, 由残差网络学习稀疏化后的图像特征, 接着对网络进行训练, 最终的神经网络可以有效地去除雨线。实验结果表明: 本文提出的方法可以有效地去除雨线。同时, 残差网络学习稀疏化的图像特征, 可以大大提升网络的训练效率。

关键词

图像特征, 计算机视觉算法, 图像去雨, 残差网络

A Single Image Derain Method Based on Improved Residual Network

Yan Cui, Ruiqi Cai*, Xiaoyi Yu

Collage of Physics and Information Engineering, Shanxi Normal University, Linfen Shanxi

Email: cuiyan8013@163.com, *270027505@qq.com, yuriiac@hotmail.com

Received: Dec. 5th, 2020; accepted: Dec. 28th, 2020; published: Jan. 5th, 2021

Abstract

Rain is a relatively common weather phenomenon outside. Under the condition of rain, the scattering and visibility of light decrease, which leads to the interference of computer vision algorithm to different degrees. In order to further reduce the influence of rain on computer vision algorithm,

*通讯作者。

this paper proposes a single image derain method based on improved residual network. This method combines the traditional image processing method and the deep learning method. Firstly, the image with rain is sparsely processed by the traditional method, and the sparse image is taken as the input. The image features after sparsity, are learned by the residual network, and then the network is trained, so that the final network can effectively remove rain lines. The experimental results show that the proposed method can effectively remove the rain line. At the same time, the residual network learning sparse image features can greatly improve the training efficiency of the network.

Keywords

Image Features, Computer Vision Algorithm, Image Remove Rain, Residual Network

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

雨水是常见的天气情况，会对室外视觉系统的性能产生挑战，雨线会使图像产生模糊效果，不同程度的降低图像视觉质量。因此在图像增强、目标跟踪等实际应用中，需要有效的去除雨线的方法。近年来，图像去雨算法被广泛研究，并且取得了良好的效果。

在过去的几十年里，学者们提出了不同的方法来降低雨线对图像质量的影响。这些方法可以分为两类：基于视频的方法和基于单图像的方法。其中基于视频的图像去雨算法利用时间的连续性可以获得冗余信息，利用这些冗余信息可以很方便的对于雨滴去除。与之相比，单幅图像去雨算法无法利用时间的连续性，因此更具有挑战性。利用字典学习稀疏编码的方法是一种常用方法[1]，首先通过滤波器将图像分为低频和高频两个部分，然后利用字典学习，把高频部分分为雨线部分和非雨线部分，最后合并保留的非雨部分和低频部分，得到去雨图像。文献[2]利用核回归和非局部均值滤波实现了雨带的检测和去除。在文献[3]中，提出了一种广义低秩模型，通过这一模型，可以实现单图像和视频降水的时空相关性学习。但是传统的方法计算时间过长，而且效果不理想。

卷积神经网络刚开始被用于手写字符的识别[4]。文献[5]论证了深度神经网络在特征提取方面的性能，从此，深度神经网络被广泛应用，至今在图像分类、行人重识别[6] [7] [8]等领域发挥着巨大的作用。为了提高网络的建模能力，文献[9]设计了一种新的尺度空间不变注意力机制，帮助网络获得部分特征。利用这种方式可以获得特征图中最活跃的的显著特征，从而获得良好的效果。文献[10]设计了一个二分支深度神经网络，来分别处理雨条纹和类雾效果。最后通过对一个子模块的联合训练得到最终的细化结果，该方法可以大大提升区域图像的视觉质量。

和传统的方法相比，虽然深度学习的方法表现更加良好，但是深度学习一般采用的神经网络结构较为复杂，卷积层数较多，需要学习的参数较多，不太适合图像去雨这种低级视觉领域任务。残差网络通过改变映射形式来简化学习过程，受到残差网络(Residual network, ResNet)的启发，本文提出一种基于改进的残差网络的单幅图像去雨算法。首先使用传统图像处理方法获得稀疏化图像，再使用残差网络提取稀疏化图像特征，通过这些特征估计出原始图像和复原图像的映射关系，最终获得复原图像。实验结果表明，本文的方法获得了较好的去雨效果。

2. 相关内容

2.1. 细节层

对于图像去雨任务，有学者把图像分为基础层和细节层，其中，基础层为图像的低频部分，细节层为图像的高频部分。无论对于有雨图像还是对应的无雨图像均成立，如公式(1)所示：

$$A = A_{\text{base}} + A_{\text{detail}} \quad (1)$$

其中， A 可以表示有雨图像或者无雨图像。 base 表示基础层， detail 表示细节层。有雨图像的基础层、细节层如图 1 所示，对应的无雨图像的基础层、细节层如图 2 所示。综合图 1 和图 2 可以看出，细节层的像素接近二值化图像，因此，神经网络学习细节层特征比较容易。有雨图像的细节层中只保留了雨线和部分物体轮廓信息，而对应的无雨图像只包含物体轮廓信息，相比于整幅图像更具稀疏性，这种稀疏性很大程度上减少了有雨图像与无雨图像之间的映射空间。网络通过直接学习细节层之间的差异，使得网络训练过程变得非常容易。

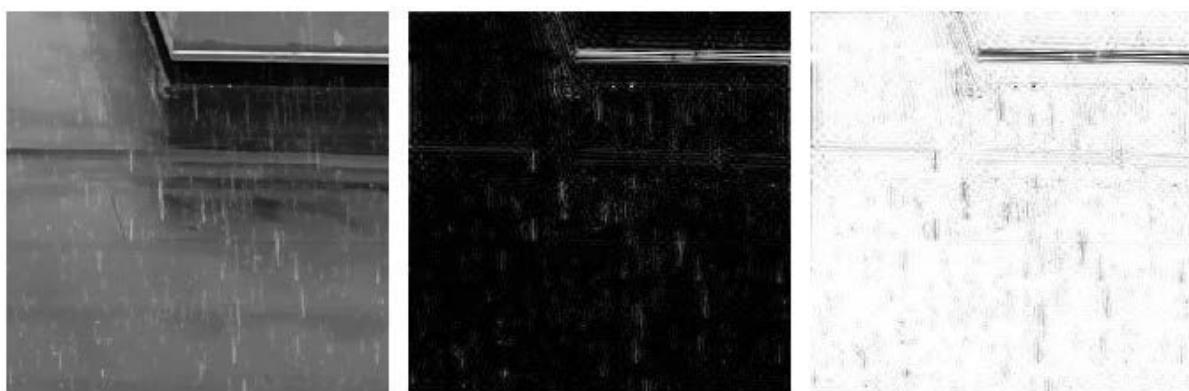


Figure 1. Rain image and its detail layer, base layer

图 1. 有雨图像及其细节层、基础层



Figure 2. No rain image and its detail layer, base layer

图 2. 无雨图像及其细节层、基础层

2.2. 负残差层

也有一些学者把有雨图像分为干净层和雨层。干净层即为对应的无雨图像。如公式(2)所示：

$$C = C_{\text{clean}} + C_{\text{rain}} \quad (2)$$

其中, clean 表示干净层, rain 表示雨层。在文献[11]中提出了负残差层, 即干净层与含雨图像相减获得的结果。负残差层如图 3 所示。

由图 3 可以观察到, 负残差层的像素值的取值范围明显缩小, 因为雨水在图像中往往以白色条纹的形式出现, 所以负残差层中雨条为白色, 雨线的背景全部变为黑色。所以, 由神经网络学习雨图像特征改为学习残差层特征, 可以大大压缩映射范围, 提高训练效率。



Figure 3. Rain image and its clean layer, negative residual layer
图 3. 有雨图像及其干净层、负残差层

3. 方法

3.1. 改进的残差网络

残差网络[12]由残差模块堆叠而成, 残差模块的结构如图 4 所示, 图像特征图作为输入, 首先经过一个 ReLU 激活函数的卷积层, 然后再次经过一个卷积层得到特征图, 给定的特征图和经过两次卷积层的特征图进行相加并通过 ReLU 得到模块的最终输出。其过程可通过公式(3)表示:

$$y_i = F(x_i, w_i) + x_i \tag{3}$$

其中 x_i 表示输入特征图, y_i 表示输出特征图, $F(x_i, w_i)$ 表示中间卷积层, w_i 表示要训练的参数。残差模块的核心思想是跳转连接的操作, 它可以解决在反向传播时梯度消失的问题, 从而为训练大型的深度网络提供了可能。

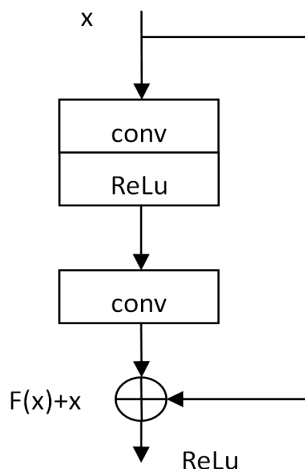


Figure 4. Residual module
图 4. 残差模块

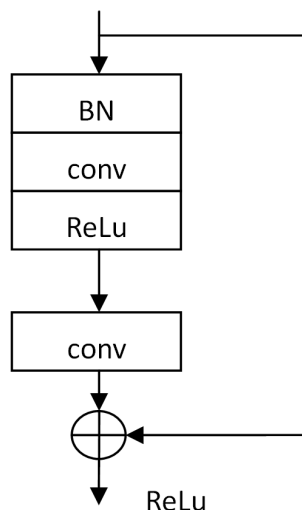


Figure 5. Improved residual module in ResNet

图 5. ResNet 中改进的残差模块

随着卷积神经网络层数的加深，在训练过程中会出现网络的收敛速度急剧下降和梯度弥散等问题。为了解决这一问题，有学者提出了利用批量归一化(BN, BatchNormalization)来解决这个问题。

批量归一化即对同一层的输入信号进行归一化，公式如下：

$$\hat{x} = \frac{x - E(x)}{\sqrt{\text{Var}(x) + \varepsilon}} \quad (4)$$

其中， x 为网络归一化的激活值， x 为网络某一层的激活值， $E(x)$ 为平均值， $\text{Var}(x)$ 为方差， ε 为极小值。BN 算法公式如下：

$$y^{(k)} = \gamma^{(k)} \hat{x}^{(k)} + \beta^{(k)} \quad (5)$$

其中，每一个神经元 $x^{(k)}$ 都会有一对这样的参数 γ 、 β 。这样其实当： $\gamma^{(k)} = \sqrt{\text{Var}[x^{(k)}]}$ ， $\beta^{(k)} = E[x^{(k)}]$ 时可以保持某一层原始的学习特征，并能重构参数 γ 、 β ，恢复初始网络学习的特征分布。

BN 层是一种规范化神经网络的激活方法，加入批量归一化算法对每一层的输入特征做归一化处理，使其数据分布相对稳定，同时在训练时设置较大的学习率，加快了网络收敛速度和训练速度。

传统残差块的排列顺序在 ResNet 中存在不足，如干路上的特征图会先通过卷积层，再到 BN 和 ReLU。输入的特征图没有先做归一化处理，所以 BN 层的存在并不能起到很大的作用。为了改善这个问题，本文把 BN 层放到了卷积层的前面，这样可以使 BN 层发挥更大的作用。改进的残差块如图 5 所示。

ResNet 网络的核心是残差块网络，本文对 ResNet 网络中的残差块网络结构进行优化调整，每个残差块包含两个卷积层，每个卷积层使用尺寸为 3×3 的卷积核来增加网络宽度，第一层卷积核前加入 BN 层是为了加快网络模型的收敛速度，提高训练后模型的准确度，第一层卷积核后加入 ReLU 防止可以提高网络的非线性表示能力，并在一定程度上防止过拟合的发生。对于整体的 ResNet 来说未调整网络深度。

输入的图像特征在干路上依次经过 BN 层、卷积层、ReLU 层、卷积层，然后在支路增加一个输入的恒等映射，使得深度网络更加快速的学习到输入到输出之间的映射关系。

3.2. 模型结构

本文通过对图像的分层，同时为了减少学习参数，提高网络运行效率，利用改进的 ResNet 网络来学

习有雨图像的细节层特征。有雨图像的细节层通过高通滤波器获得，这里选取的高通滤波器为双边滤波器。ResNet 网络的核心就是残差网络，通过对残差网络的结构进行改进，从而进一步提升 ResNet 网络学习特征的能力。网络结构如图 6 所示。



Figure 6. Image derain network structure

图 6. 图像去雨网络结构

本文提出的模型输入为有雨图像，有雨图像通过双边滤波器获得细节层，负残差层需要经过图像预处理获得，即用有雨图像减去无雨图像。并通过改进的 ResNet 网络来学习细节层到负残差层的映射，最后将输入和学习到的负残差层相加，得到的输出即为去雨图像。

3.3. 损失函数

本文使用去雨图像与原始图像的均方误差作为损失函数，公式如下所示：

$$L(P) = \frac{1}{n} \sum_{i=1}^n \|h_p(I_i) - J_i\|^2 \quad (6)$$

其中， $h_p(I_i)$ 为去雨图像， J_i 为原始图像，利用随机梯度下降的方法使均方误差最小化，得到参数的最优值。

3.4. 训练方法

本文的训练和测试所需数据集为文献[11]提供的数据集。该数据集有 1000 对图片，每对图片由一张无雨图像和对应的十四张有雨图像组成，有雨图像通过 PhotoShop 人工添加雨线形成。该实验随机选取 800 对图片训练神经网络，剩下的 200 对图片用于测试网络模型。模型的仿真实验是在带有 NVIDIA Tesla P100GPU 的机器上进行的。批训练的大小设为 2，利用 ADAM 优化器来训练模型，并且将 η 设置为 0.001， β_1 设置为 0.9， β_2 设置为 0.999。初始化学率为 0.0001，每训练 100 次后学习率降低 10 倍，模型训练 500 个 epoch。

3.5. 实验分析

在图 7 中展示出了直观的视觉效果图像。从左往右，依次是有雨图像、DDN [13]、Pix2Pix [14] 方法的预测图以及本文方法的结果，从图像上可以直观感受到采用我们的方法可以去除大部分雨线，并且还还原了背景图像的部分颜色和纹理细节。虽然这三种方法无法使得雨线去除干净，但是本文模型相较于 DDN 方法有更加少地雨线，相较于 Pix2Pix 方法有更加清晰地背景图像纹理，说明我们的网络模型表现良好。

在表 1 中展示了本文提出的方法与 DDN、Pix2Pix 的性能比较。采用峰值信噪比 (PSNR) 和结构相似性 (SSIM) 作为数值评价指标。峰值信噪比可以反映图像质量的好坏，结构相似性可以反映去雨图像和无雨图像之间的亮度、对比度和结构的相似性。其中，第一、二行数据为 DDN、Pix2Pix 方法、本文方法对训练集中随机两张有雨图像的处理结果，由于随机图像的结构不同，结果可能会产生很大的随机性，因此第三行为 DDN、Pix2Pix、本文方法对整个测试集的处理平均结果。

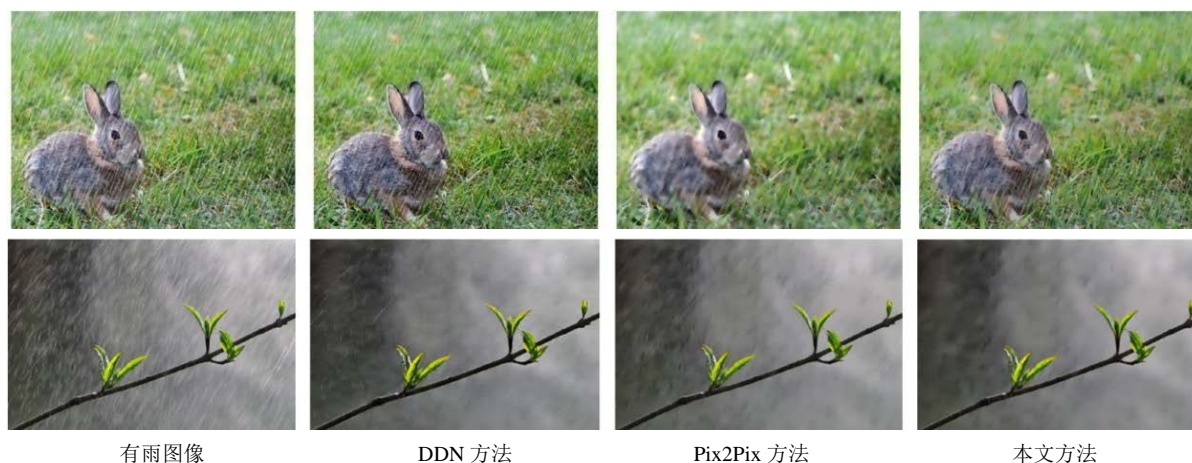


Figure 7. Experimental results of different methods

图 7. 不同方法的实验结果

从表 1 中可以看出,对于数据集中的随机两张有雨图像,本文方法的 PSNR 和 SSIM 优于 DDN 方法,在 PSNR 方面优于 Pix2Pix 方法。对于整个测试集的处理平均结果,DDN 方法所得结果的峰值信噪比和结构相似性分别为 28.04 和 0.82, Pix2Pix 方法所得结果的峰值信噪比和结构相似性分别为 32.65 和 0.94, 本文方法所得结果的峰值信噪比和结构相似性分别为 34.14 和 0.88, 在测试集中我们的方法虽然在 SSIM 上的表现不如 Pix2Pix 模型,但是优于 DDN 的表现,并且在 PSNR 上比 DDN 和 Pix2Pix 表现更好,即图像质量更高。说明通过改进的 ResNet 网络学习图像特征,可以获得良好的细节层到负残差层之间的映射关系。

Table 1. Performance comparison of different methods

表 1. 不同方法的性能比较

图像	峰值信噪比		
	DDN	Pix2Pix	本文方法
1	26.67	33.06	35.71
2	27.52	35.81	36.26
200 对测试图像	28.04	32.65	34.14
图像	图像相似性		
	DDN	Pix2Pix	本文方法
1	0.73	0.94	0.86
2	0.85	0.96	0.90
200 对测试图像	0.82	0.94	0.88

4. 结论

本文提出了基于改进的残差网络的单幅图像去雨方法,通过调整 ResNet 中的残差结构,让残差结构更适合图像去雨这种低视觉领域任务,使得非线性映射关系估计更为准确,提高了复原图像质量;通过对图像稀疏化,使得图像特征学习更加简单。本文算法的网络结构仍然有不足之处,例如网络中残差块的数量不同会导致实验结果的差异,需要进一步研究,以提高图像去雨效果。

参考文献

- [1] Kang, L.W., Lin, C.W. and Fu, Y.H. (2011) Automatic Single-Image-Based Rain Streaks Removal via Image Decomposition. *IEEE Transactions on Image Processing*, **21**, 1742-1755. <https://doi.org/10.1109/TIP.2011.2179057>
- [2] Kim, J.H., Lee, C., Sim, J.Y., et al. (2013) Single-Image Deraining Using an Adaptive Nonlocal Means Filter. 2013 *IEEE International Conference on Image Processing*, Melbourne, 15-18 September 2013, 914-917. <https://doi.org/10.1109/ICIP.2013.6738189>
- [3] Chen, Y.L. and Hsu, C.T. (2013) A Generalized Low-Rank Appearance Model for Spatio-Temporally Correlated Rain Streaks. *Proceedings of the IEEE International Conference on Computer Vision*, Sydney, 1-8 December 2013, 1968-1975. <https://doi.org/10.1109/ICCV.2013.247>
- [4] Pang, B., Zhai, D., Jiang, J., et al. (2020) Single Image Deraining via Scale-Space Invariant Attention Neural Network. *Proceedings of the 28th ACM International Conference on Multimedia*, 375-383. <https://doi.org/10.1145/3394171.3413554>
- [5] Wang, Y., Gong, D., Yang, J., et al. (2019) An Effective Two-Branch Model-Based Deep Network for Single Image Deraining. *Computer Vision and Pattern Recognition*, arXiv: 1905.05404 [cs.CV].
- [6] Tan, M. and Le, Q.V. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *International Conference on Machine Learning*, arXiv: 1905.11946 [cs.LG].
- [7] Zoph, B., Vasudevan, V., Shlens, J., et al. (2018) Learning Transferable Architectures for Scalable Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 8697-8710. <https://doi.org/10.1109/CVPR.2018.00907>
- [8] Zheng, M., Karanam, S., Wu, Z., et al. (2019) Re-Identification with Consistent Attentive Siamese Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 5728-5737. <https://doi.org/10.1109/CVPR.2019.00588>
- [9] Li, X., Wu, J., Lin, Z., et al. (2018) Recurrent Squeeze-and-Excitation Context Aggregation Net for Single Image Deraining. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds., *Computer Vision-ECCV 2018, ECCV 2018, Lecture Notes in Computer Science*, Springer, Cham, 262-277. https://doi.org/10.1007/978-3-030-01234-2_16
- [10] Ge, Y., Chen, D. and Li, H. (2020) Mutual Mean-Teaching: Pseudo Label Refinery for Unsupervised Domain Adaptation on Person Re-Identification. *Computer Vision and Pattern Recognition*, arXiv: 2001.01526 [cs.CV].
- [11] Fu, X., Huang, J., Zeng, D., et al. (2017) Removing Rain from Single Images via a Deep Detail Network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 1715-1723. <https://doi.org/10.1109/CVPR.2017.186>
- [12] He, K., Zhang, X., Ren, S., et al. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [13] Fu, X., Huang, J., Ding, X., et al. (2017) Clearing the Skies: A Deep Network Architecture for Single-Image Rain Removal. *IEEE Transactions on Image Processing*, **26**, 2944-2956. <https://doi.org/10.1109/TIP.2017.2691802>
- [14] Isola, P., Zhu, J.Y., Zhou, T., et al. (2017) Image-to-Image Translation with Conditional Adversarial Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 5967-5976. <https://doi.org/10.1109/CVPR.2017.632>