Single-image-based Rain Detection and Removal via CNN

Tianyi Chen¹,ₐ and Chengzhou Fu¹,ᵇ
¹School of Computer, South China Normal University, Guangzhou 510631, China
*Corresponding author. ᵃtychen@m.scnu.edu.cn, ᵇfucz@m.scnu.edu.cn

Abstract. The quality of the image is degraded by rain streaks, which have negative impact when we extract image features for many visual tasks, such as feature extraction for classification and recognition, tracking, surveillance and autonomous navigation. Hence, it is necessary to detect and remove rain streaks from single images, which is a challenging problem since we have no spatial-temporal information of rain streaks compared to the dynamic video stream. Inspired by the priori that the rain streaks have almost the same feature, such as the direction or the thickness, although they are in different types of real-world images. The paper aims at proposing an effective convolutional neural network (CNN) to detect and remove rain streaks from single image. Two models of synthesized rainy image, linear additive composite model (LACM model) and screen blend model (SCM model), are considered in this paper. The main idea is that it is easier for our CNN network to find the mapping between the rainy image and rain streaks than between the rainy image and clean image. The reason is that rain streaks have fixed features, but clean images have various features. The experiments show that the designed CNN network outperforms state-of-the-art approaches on both synthesized and real-world images, which indicates the effectiveness of our proposed framework.

1. Introduction
Clean image is necessary in computer vision research and application, such as feature extraction for classification or recognition, tracking, surveillance and autonomous navigation. Under rainy conditions, the qualities of image are often degraded by rain streaks. Effective methods to remove rain streaks from single image are becoming research hot in recent years. Single-image-based method and video-based method are two main tasks for rain detection and removal. Compared to video-based method, it is more difficult to do rain detection and removal by single-image-based method due to the following reasons.

- There are no temporal contents of rain streak in single image.
- Some details of the image are similar to the rain streak, which makes it hard to detect and remove the rain streak and keep the original image structure wholly.

In this paper, we propose an effective CNN network to detect the rain streak in the rainy image, thus by subtracting the rain streak, we can obtain the original clean image.

2. Background
For the rain streak in the dynamic video stream, it is easy to identify the rain streak by comparing the differences of the continuous video frames [1]. To detect and remove rain streak in single image is more challenging. J. H. Kim et al. [2] proposed a mean filtering and a kernel regression to detect and remove rain streaks. The patch-based approaches [3, 4] are used to remove rain streaks effectively. In recent years, the CNN [5] has been designed to do rain-removal and the satisfactory results have been gotten. Based on the deep CNN, the mapping from rainy image to clean image detail layer is learnt. We know that the mapping depends on each pair of the rainy image and clean image, which results in less robust of the CNN parameters. We notice that the rain streaks in different rainy images have the same features, such as the direction or the thickness. In this paper, we design a new CNN to get the
mapping between the rainy images and the rain streaks. Thus, the clean image is obtained by subtracting the rain streaks from the rainy images.

In this paper, we consider two models of synthesized rainy image, LACM model and SCM model.

- LACM model [1,3,5,6]
  \[ I = B + R, \]  

- SCM model [4]
  \[ I = B + R - B \cdot R, \]  

Where \( I \) denotes a rainy image, which is composited by two layers, de-rained image layer \( B \) and rain layer \( R \). Here “\(-\)” denotes point-wise multiplication operator. Some examples are shown in Figure 1.

![Figure 1](image)

(a) Clean image (b) LACM (c) SCM

Figure 1 Some examples of clean image, LACM-model image and SCM-model image.

By guided image filter [7], a rainy image \( I \) could be decomposed to obtain the low-pass layer \( I_{\text{base}} \) and the high-pass layer \( I_{\text{detail}} \),

\[ I = I_{\text{detail}} + I_{\text{base}}, \]  

Where \( I_{\text{base}} \) and \( I_{\text{detail}} \) contain the smooth part and the detail part of \( I \) respectively. It is obvious that the rain layer \( R \) is high-pass, which is included in \( I_{\text{detail}} \). Thus, the high-pass layer \( I_{\text{detail}} \) is mainly studied as shown in [5], which makes it easier to find the suitable mapping between \( I_{\text{detail}} \) and the rain streaks. Moreover, the computing cost is lower when only the \( I_{\text{detail}} \) is considered. A CNN network is trained in [5] instead of the raw rainy image to obtain the clean image.

Although the performance is better in [5] to remove the rain streaks from the rainy image. We know that the CNN network [5] is to directly find the mapping between \( I_{\text{detail}} \) and the clean image. There are so many kinds of rainy images, which make the CNN network unstable since it depends on each pair of the \( I_{\text{detail}} \) and the clean image. Thus, the effect is not robust when the rainy images, which are not like the training images, are tested. We notice that the rain streaks in different types of rainy images almost have the same features, such as the direction or the thickness of the rain streaks. In this paper, a new CNN network is proposed to find the mapping between the rainy images and the rain streaks, instead of the mapping between the rainy images and the clean images. The rain streaks are taken as the supervised object in our CNN network whose destination is to learn the features of the rain streaks. So, the result does not depend on each pair of the input images. Moreover, compared to the complex features of various clean images, the features of the rain streaks are more stable, which makes our CNN network more robust. Thus, our CNN network is suitable to different kinds of rainy images, including synthesized and real-world rainy images.

3. Our proposed method

3.1. Customized-filter rainy detection and removal network (CRDRN)

We propose a novel rain detection and removal model. The input images are the rainy image \( I \) and the rain streak \( R \). A customized-filter rain detection and removal network(CRDRN) is designed to detect and remove the rain streak \( R \) in \( I \), which minimizes the objective function as follows:
\[ L = \sum_{n=1}^{N} || f(l^n_{\text{detail}}, W, b) - R^n ||^2, \quad (4) \]

Where \( N \) is the number of training images and \( n \) indexes the image, \( f(\cdot) \) is CRDRN, \( W \) and \( b \) are network parameters that need to be learned, \( l^n \) is the detail layer obtained by guided image filter, \( R \) is the supervised rain streaks. Our CRDRN network architecture is shown in Figure 2.

**Figure 2** The proposed CRDRN network architecture for single-image-based rain detection and removal.

### 3.1.1. Rain detection and removal

Our CRDRN network consists of three CNN layers to detect and remove rain streaks, which can be expressed as the following operations:

\[
\begin{align*}
    I_{\text{detail}} &= I - I_{\text{base}}, \\
    l^1 &= \sigma(W^1 * I_{\text{detail}} + b^1), \quad l = 1, \\
    l^2 &= \sigma(W^2 * l^{1-1} + b^2), \quad l = 2, \\
    l^3 &= \sigma(W^3 * l^{2-1} + b^3), \quad l = 3, \\
    O &= I_{\text{detail}} - l^1 + l_{\text{base}}, \quad l = 3,
\end{align*}
\]

Where \( l = 1, 2, 3 \) indexes the layer number, “\( * \)” is the convolution operation, \( W_l \) is the trained weight and \( b_l \) is the bias. The \( \sigma(\cdot) \) is ReLU nonlinear activation function [8]. The ReLU is effective to truncate the parameters less than 0, which makes the CRDRN network parameters take normal values. Two hidden layers are used in our CRDRN network and \( O \) is the output.

For the first layer in CRDRN, we use filters of size \( s_1 \times s_1 \times b \times c \) to obtain \( c \) feature maps by \( b \) input image channels; \( s \) denotes filter size. For the second layer, the size of filters is \( s_2 \times s_2 \times c \times c \), which extract the rain streaks from the rainy images. For the last layer, the size of filters is \( s_3 \times s_3 \times c \times b \), which turn \( c \) feature map into \( b \)-dimension output image.

### 3.1.2. Training

We use Adam optimizer [9] and take the average of 50000 iterated times to minimize the objective function (4). We selected 350 clean images to generate training set as in [5]. Based on two generation models (LACM model and SCM model) and seven different streak orientations, each clean image was used to synthesize 14 pairs of rainy images and rain streaks. Thus, a dataset containing \( 350 \times 14 = 4900 \) pairs of rainy images and rain streaks is set up. We randomly selected 40 \( 64 \times 64 \times 3 \) rainy image/rain streak patch pairs from each pair. The final training dataset contains 196000 rainy image/rain streak patch pairs.
4. Experiment

4.1. Parameters setting
To evaluate the performance of our CRDRN, both synthesized and real-world rainy images are tested. Three recent high-quality methods [3,4,5] are compared with ours. Software implementations of our method were provided by Python based on Tensorflow. All experiments are carried on a personal computer with Intel Core CPU i7-7700, 32GB RAM and NVIDIA GeForce GTX 1080. Our CRDRN network consists of two hidden layers and one output layer. The kernel sizes are set as $s_1 = 5$, $s_2 = 1$ and $s_3 = 3$ respectively. The numbers of convolution kernels are $b = 3$ and $c = 512$. All these parameters $s$, $b$ and $c$ are defined in Sec. 3.1.1. The learning rate of Adam optimizer is set as $10^{-4}$, and the mini-batch size is set as 50.

4.2. Results on synthesized rainy image
Firstly, we evaluate the testing results on synthesized rainy images. Rain12 [3], which consists of 12 synthesized rainy images and is a benchmark dataset for rain removal, is selected as the testing dataset. The synthesized rainy images in Rain12 are generated using photorealistic rendering techniques [10].

![Figure 3. Example results on synthesized rainy images “cat” and “temple” respectively, which were for testing and not for training. The odd row shows the full image and the even row shows one zoomed-in region.](image)

Figure 3 shows visual comparisons for two synthesized rainy images “cat” and “temple” respectively. As can be seen, the outputs of method [4] still contain many short rain streaks although they are lighter. Method [5] is better, but some rain streaks are remained in the images. However, as shown in the last column, our CRDRN method can detect and remove rain streaks.

Since we have the ground truth, the SSIM [11] could be used for quantitative evaluation. A higher SSIM value means a better result in terms of image structural properties. As shown in Tab.1, our method has the highest SSIM values, which is in agreement with the visual results in Figure 3.

| Pic1 | Pic2 | Pic3 | Pic4 | Pic5 | Pic6 | Pic7 | Pic8 | Pic9 | Pic10 | Pic11 | Pic12 | Average |
|------|------|------|------|------|------|------|------|------|-------|-------|-------|---------|
| DSC  | 0.86 | 0.92 | 0.86 | 0.98 | 0.93 | 0.97 | 0.97 | 0.91 | 0.94  | 0.86  | 0.90  | 0.87    | 0.91    |

Table 1. Each quantitative comparison results on Rain12 [3] benchmark dataset using SSIM.
Moreover, we selected 100 images, which are never used in training set, from BSD300 [12] dataset. LACM model is used to generate rainy images as testing dataset with random directions. SSIM is used for quantitative evaluation. As shown in Tab.2, our method has the highest SSIM values compared to DSC method [4] and DerainNet method [5].

Table 2. Quantitative measurement results on BSD300 [12] using SSIM.

|       | DSC | DerainNet | Ours |
|-------|-----|-----------|------|
| SSIM  | 0.92| 0.93      | 0.94 |

Figure 4. Example results on real-world rainy image “soccer” and “street”.

4.3. Results on real-world rainy image

For we have no ground truth for real-world rainy images, our CRDRN method is tested on real-world images using the network trained on the 196000 rainy image/rain streak patch pairs as described in Sec. 3. As shown in Figure 4, the outputs of method [4] still contain many rain streaks. There are less rain streaks in the images of method [5], but the images are blurred. However, as shown in the last column, our CRDRN method not only detects and removes rain streaks, but also make the image clearer, which indicates that our CRDRN method behaves well for rain detection and removal for real-world rainy image.

5. Conclusion

In this paper, we propose a CNN network architecture called CRDRN for rain detection and removal from single images. Based on CRDRN network our method learns the features of rain streaks of rainy
image detail layers. Although the CRDRN network is trained on synthesized clean/rainy image pairs, it shows better performances on both synthesized dataset and real-world dataset, which indicates that our CRDRN network is effective and robust for rain detection and removal. The experiments show that our CRDRN method outperforms other state-of-the-art methods regarding image quality.

In future, we will work on setting up an image dataset for better evaluating framework for single-image rain detection and removal. In addition, we will also investigate the applicability of the proposed method to other challenges in computer vision, for example, single-image snow detection and removal.

Acknowledgement
Tianyi Chen thanks the supports by the Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation ( “Climbing Program 2018” Special Funds, No.pdjha0129), and the Project of Golden Seed Cultivation (No.17JJKC06). Chengzhou Fu thanks the supports by Guangdong Science and Technology Project (No.2013B090800024), the Project of Young Education Research and Cultivation Fund of South China Normal University (No.17KJ16), and the Science and Technology Program of Guangzhou, China (No.201604016007).

References
[1] Barnum P C, Narasimhan S and Kanade T 2010 Analysis of rain and snow in frequency space Int. Journal of Computer Vision 86 256–274
[2] Kim J H, Lee C, Sim J Y and Kim C S 2013 Single-image deraining using an adaptive nonlocal means filter Int. Conf. on Image Processing 914-917
[3] Li Y, Tan R T, Guo X, Lu J and Brown M S 2016 Rain streak removal using layer priors IEEE Conf. on Computer Vision and Pattern Recognition 2736-2744
[4] Luo Y, Xu Y and Ji H 2015 Removing rain from a single image via discriminative sparse coding Int. Conf. on Computer Vision 3397-3405
[5] Fu X, Huang J, Ding X, Liao Y, Paisley J 2016 Clearing the skies: a deep network architecture for single-image rain streaks removal IEEE Transactions on Image Processing 26 2944-2956
[6] Kang L W, Lin C W and Fu Y H 2012 Automatic single-image-based rain streaks removal via image decomposition IEEE Transactions on Image Processing 21 1742–1755
[7] He K, Sun J and Tang X 2013 Guided image filtering IEEE Transactions on Pattern Analysis and Machine Intelligence 35 1397-1409
[8] Krizhevsky, Alex, Sutskever I and Hinton G E 2012 ImageNet classification with deep convolutional neural networks Annual Conf. on Neural Information Processing Systems 25 1097-1105
[9] Kingma D P and Ba J 2014 Adam: A method for stochastic optimization Preprint arXiv:1412.6980
[10] Garg K and Nayar S K 2006 Photorealistic rendering of rain streaks ACM Transactions on Graphics 25 996–1002
[11] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: From error visibility to structural similarity IEEE Transactions on Image Processing 13 600–612
[12] Martin D, Fowlkes C and Tal Dand Malik J 2001 A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics Int. Conf. on Computer Vision 2 416–423