Detail-recovery Image Deraining via Context Aggregation Networks

Sen Deng[1,2] Mingqiang Wei[1,2*] Jun Wang[1,2] Yidan Feng[1,2]
Luming Liang[3] Haoran Xie[4*] Fu Lee Wang[5] Meng Wang[6]
[1]Nanjing University of Aeronautics and Astronautics
[2]MIIT Key Laboratory of Pattern Analysis and Machine Intelligence
[3]Microsoft Applied Sciences Group [4]Lingnan University
[5]The Open University of Hong Kong [6]Hefei University of Technology

Abstract

This paper looks at this intriguing question: are single images with their details lost during deraining, reversible to their artifact-free status? We propose an end-to-end detail-recovery image deraining network (termed a DRD-Net) to solve the problem. Unlike existing image deraining approaches that attempt to meet the conflicting goal of simultaneously deraining and preserving details in a unified framework, we propose to view rain removal and detail recovery as two separate tasks, so that each part could be specialized rather than traded off. Specifically, we introduce two parallel sub-networks with a comprehensive loss function which synergize to derain and recover the lost details caused by deraining. For complete rain removal, we present a rain residual network with the squeeze-and-excitation (SE) operation to remove rain streaks from the rainy images. For detail recovery, we construct a specialized detail repair network consisting of well-designed blocks, named structure detail context aggregation block (SDCAB), to encourage the lost details to return for eliminating image degradations. Moreover, the detail recovery branch of our proposed detail repair framework is detachable and can be incorporated into existing deraining methods to boost their performances. DRD-Net has been validated on several well-known benchmark datasets in terms of deraining robustness and detail accuracy. Comparisons show clear visual and numerical improvements of our method over the state-of-the-arts[1].

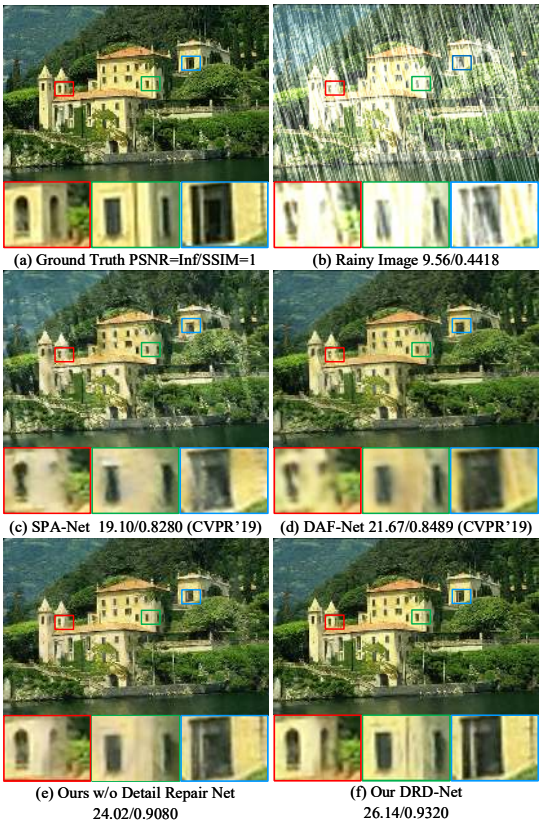


Figure 1. Image deraining results tested in the dataset of Rain200H. From (a)-(f): (a) the ground truth image Castle, (b) the rainy image Castle and the deraining results of (c) SPA-Net [25], (d) DAF-Net [10], (e) Ours w/o Detail Repair Net, and (f) our DRD-Net respectively.

1. Introduction

Images captured in rainy days inevitably suffer from noticeable degradation of visual quality. The degradation causes detrimental impacts on outdoor vision-based tasks, such as video surveillance, autonomous driving, and object detection. It is, therefore, indispensable to remove rain in rainy images, which is referred to as image deraining.

The ultimate goal of image deraining is to recover the ground-truth image $\mathbf{B}$ from its observation $\mathbf{O} = \mathbf{B} + \mathbf{R}$ with the rain streaks $\mathbf{R}$, which is an ill-posed problem since

*Co-corresponding authors (mqwei@nuaa.edu.cn/hrxie@ln.edu.hk).
[1]Source code: https://github.com/Dengsgithub/DRD-Net

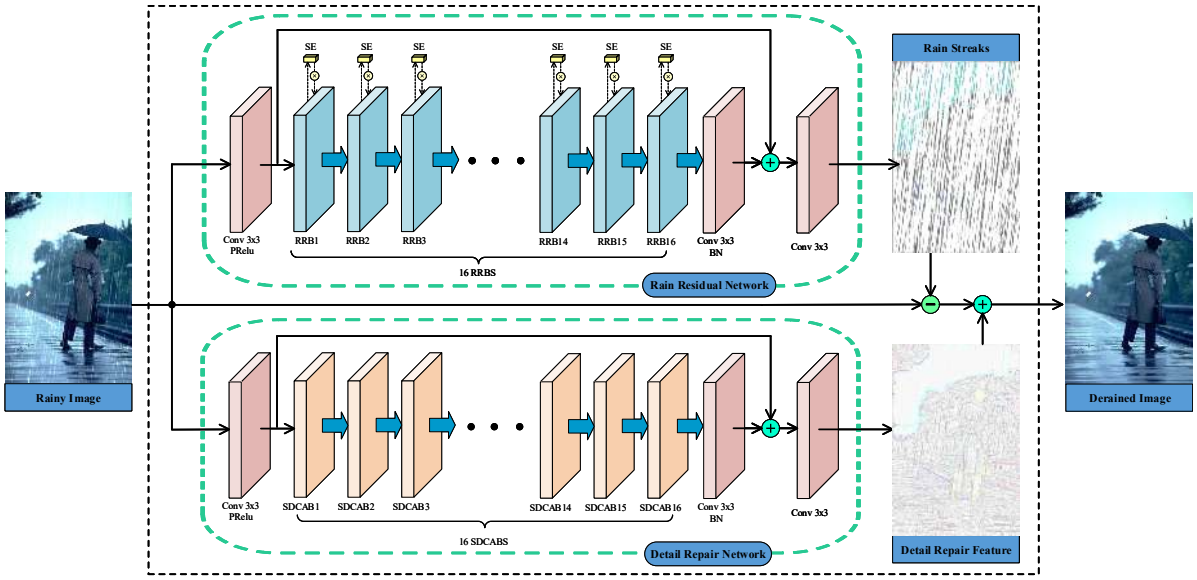


Figure 2. DRD-Net consists of two-sub networks, i.e., the rain removal network and the detail repair network. The first sub-network, which combines the squeeze-and-excitation (SE) operation with residual blocks to make full advantage of spatial contextual information, aims at removing rain streaks from the rainy image. And the second sub-network, which integrates the structure detail context aggregation block (SDCAB) to aggregate context feature information for a large reception field, seeks to recover the lost details to the derained image.

both the clean image and rain streaks are all unknown. The traditional image deraining approaches behave ineffectively in medium- and heavy-rain removal, while the learning-based ones commonly lead to image degradations such as the loss of image details, halo artifacts and/or color distortion. Video-based deraining methods could borrow the redundant information between the sequence frames for quality rain removal [6, 22]. In contrast, single image based deraining methods ought to either draw support from priors, such as Gaussian mixture model [17], sparse coding [19] and low-rank representation [29, 7] or feed a large dataset into the well-designed deep networks [5, 10, 25], due to the lack of sequence information.

Despite the great improvements of image deraining to produce promising deraining results when handling light-rain images, they are hindered to both remove rain streaks completely and preserve image details effectively on the images captured from the extremely bad weather. Such a phenomenon happens in Fig. 1. That is because the magnitude of image details is similar to and even smaller than that of rain streaks, but a rainy image in computer lacks semantic information to describe them separately. Therefore, the image details and rain streaks are commonly removed simultaneously. No state-of-the-art methods can serve as an image deraining panacea for various applications: they produce the deraining results with a tradeoff between rain removal and image detail maintenance.

Different from existing image deraining methods that attempt to maintain image details and lack the detail-recovery mechanism, we look at this intriguing question: now that image deraining leads to image detail blurring in nature, are single images with their details lost during image deraining, reversible to their artifact-free status?

We propose an end-to-end detail-recovery image deraining network (DRD-Net) based on the context aggregation networks, which simply introduces a collateral branch but succeed in recovering the original image details clearly. The main contribution of this work can be concluded as follows:

- We build a two-branch parallel network (DRD-Net) composed of a squeeze-and excitation (SE) based rain residual network and a detail repair network. SE aggregates feature maps in the same convolutional layer to make full advantage of spatial contextual information for complete rain removal and the additional detail repair network encourages the lost details to return to the image after deraining by the rain residual network.

- We present the *structure detail context aggregation block* (SDCAB), which has larger reception fields and makes full use of the rain-free image patches, and demonstrate how SDCAB could facilitate the specific task of detail recovery.

- Our proposed framework regards rain removal and detail recovery as two independent tasks, therefore the detail recovery branch is actually detachable and can be incorporated into existing deraining methods to improve their performances.

2. Related Work

2.1. Video Deraining Methods

Owing to the redundant information of the sequence frames in videos, rain streaks can be more easily identified and removed [6, 22, 1]. [6] replaces the intensity of a rainy pixel by averaging the corresponding pixel's intensities in the adjacent frames. [2] detects the rain streaks based on the histogram of rain streaks orientations. [23] summarizes the video-based deraining methods that have been proposed in recent years. In addition, the deraining results from video-based techniques may serve as the clean images for single image deraining [25].

2.2. Single Image Deraining Methods

Without the temporal information, single image-based methods are more challenging than video-based methods. For rain removal from single images, existing methods fall into two categories: the traditional methods and the deep-learning based methods.

Traditional Methods: Various image priors have been proposed to remove rain from single images. They assume that rain streaks **R** are sparse and in similar directions. Under this assumption, they decompose the input image **O** into the rain-free background scene **B** and the rain streaks layer **R**. [12] separates the rain streaks from high frequency using dictionary learning. [19] presents a discriminative sparse coding for separating rain streaks from the background image based on image patches. In [17], Gaussian mixture models (GMM), as a prior, is proposed to decompose the input image into the rain streaks and the background layer. [33] first detects rain-dominant regions and then the detected regions are utilized as a guidance image to help separate rain streaks from the background layer. [29] leverages the low-rank property of rain streaks to separate the two layers.

Deep Learning-based Methods: Deep-learning based methods have been introduced to single image deraining by [5], which boost the deraining performance significantly. Later, [31] presents a conditional generative adversarial network (GAN) and uses the perceptual loss to refine the results. [27] develops a deep recurrent dilated joint rain streaks detection and removal network to remove the rain streaks. [16] proposes the multi-stage networks based on the recurrent neural network architecture to remove rain streaks in different directions. [30] presents a density-aware multi-stream connected network for deraining. By maintaining negative residual features, [4] builds a residual-guided network for removing the rain streaks from single images and [18] exploits the potential of paired operations via dual residual connection. Unlike [27] which regards rain accumulation removal as a separate task, [10] and [15] merge physics formulation into the whole network architecture. To handle real-world cases, [25] constructs a real-world rain dataset and incorporates the spatial attentive mechanism into the network design while [26] considers both supervised images pairs and unsupervised real rainy images into the network training by minimizing the KL distance between parameterized distributions of their rain residuals. In [24], the encoder-decoder based network is interpreted as a conditional generator and the deraining performance is improved by a residual learning branch optimizing the input of the generator.

Existing image deraining approaches attempt to maintain image details by using appropriate loss functions or learning from a large even real-world dataset. However, image details are still lost to a certain extent, due to many complicated reasons. For example, most of current network architectures cannot deal with the saturated regions where the rain is very heavy and totally occlude the background scenes. That is the reason why a detail repair network becomes necessary. Thus, we propose a detail-recovery image deraining network, which comprises of two sub-networks with a comprehensive loss function for synergizing to derain and recover the lost details caused by deraining.

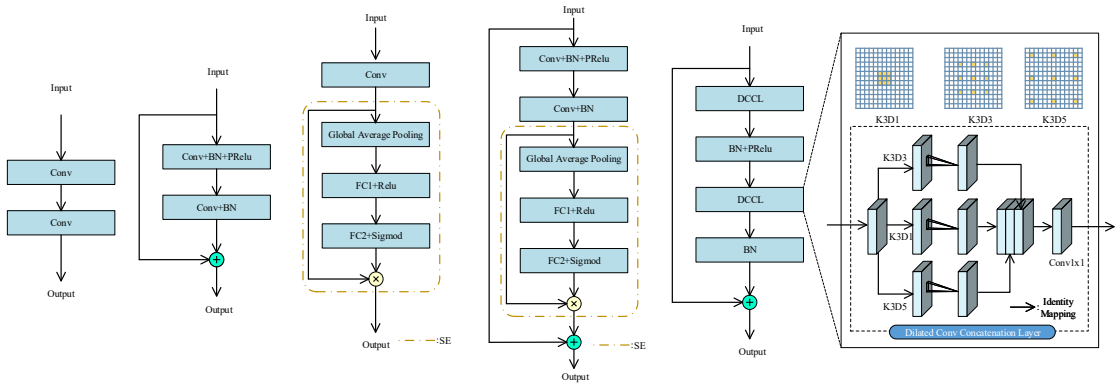
3. DRD-Net

Image deraining usually leads to detail blurring, because rain streaks and image details are all of high frequency in nature and they inevitably share similar geometrical properties. Unfortunately, existing approaches pay little attention to recovering the image details once they are lost during image deraining. For both rain removal and detail recovery of single images, we propose two sub-networks which work together as shown in Fig. 2. On one hand, we introduce a rain residual network to train a function that maps the rainy images to their rain streaks. Therefore, we can obtain the preliminarily derained images by separating the rain streaks from the rainy images. On the other hand, different from other methods which try to decompose a single rainy image into a background layer and a rain streaks layer, we present an additional detail repair network to find back the lost details. In the following, we will introduce the rain residual network and the detail repair network, respectively.

3.1. Rain Residual Network

Residual learning is proved to be a powerful tool for image restoration tasks like denoising and deraining [32, 16]. Based on the observation that the rain streaks $\mathbf{R}$ are sparser than the rain-free background scene $\mathbf{B}$ [16], we learn a function by training a residual network, which maps the rainy image $\mathbf{O}$ to rain streaks $\mathbf{R}$. We train such a network by minimizing the loss function as

$$Loss_r = \sum_{i \in N(D)} ||f(\mathbf{O}_i) - \hat{\mathbf{R}}_i||^2, \tag{1}$$



(a) Direct Network (b) Residual Block (c) Direct Network with SE (d) Residual Block with SE (d) Structure Detail Context Aggregation Block

Figure 3. Different convolution styles. From (a)-(d): (a) direct network, (b) residual block, (c) direct network with SE [16], (d) rain residual block with SE used in our rain residual network, and (e) structure detail context aggregation block used in our detail repair network.

where the $f(\cdot)$ is a function that we try to learn, $\mathbf{O}_i$ is a rainy image and $\hat{\mathbf{R}}_\mathbf{i}$ is the ground-truth rain streak layer in the training dataset D whose number is $N(D)$.

The architecture of our rain residual network is shown in the upper part of Fig. 2, which utilizes the Squeeze-and-Excitation (SE) [9] operation. Considering that the skip-conections can provide long-range information compensation and enable the residual learning [14], we combine the SE operation with the residual block in our rain residual network, which is different from Fig. 3(c) used in RESCAN [16]. The rain residual network includes 3 convolution layers and 16 rain residual blocks. The first layer can be interpreted as an encoder, which is used to transform the rainy image into the feature maps, and the last two layers are used to recover the RGB channels from feature maps.

Mathematically, the rain residual block can be formulated as

$$RRB = SE(Res(\mathbf{X_0})), \tag{2}$$

where RRB is the output of the rain residual block, $SE(\cdot)$ and $Res(\cdot)$ denote the SE operation and the residual block shown in Fig. 3(d) respectively, and $\mathbf{X_0}$ is the input signal.

Spatial contextual information has proved to be effective in single image deraining [11, 16]. Nevertheless, the different feature channels in the same layer are independent and have little correlation during the previous convolution operation. A main difference from the common residual block is that we combine SE into the residual block in our network. Since SE can model a correlation between different feature channels, we can intensify the feature channel which has more context information by giving a larger weight. Conversely, the feature channels that have less spatial contextual information will just receive a small weight. All the weights of different channels are learned by the rain residual network automatically during the training steps.

3.2. Detail Repair Network

Now that image deraining leads to image degradations in nature, we can train additional detail-recovery network that makes the detail-lost images be reversible to their artifact-free status. Based on the preliminarily derained image $\mathbf{I}_p$ which is obtained by subtracting the rain streaks $\mathbf{R}$ from the rainy image $\mathbf{O}$, we can train a function to encourage the lost details to return by optimizing the loss function as

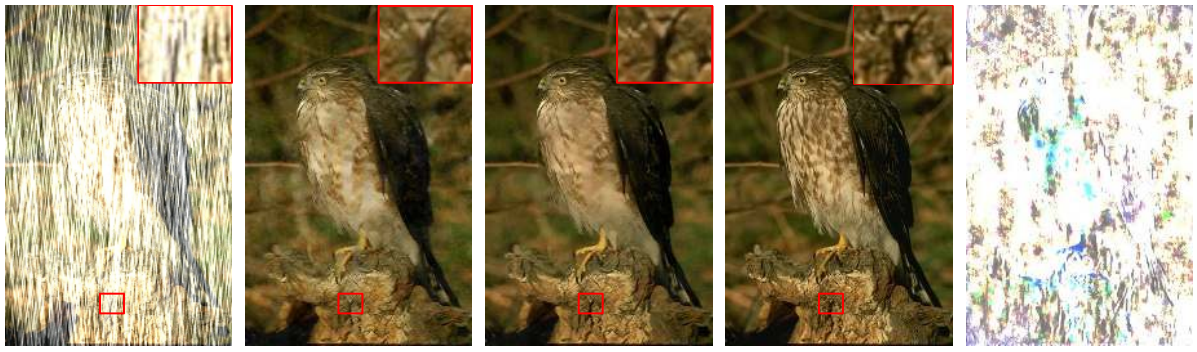
$$Loss_d = \sum_{i \in N(D)} ||(\mathbf{I}_{p,i} + g(\mathbf{O}_i)) - \hat{\mathbf{I}}_i||^2, \tag{3}$$

where $g(\cdot)$ is a function that we try to learn, $\mathbf{O}_i$ is a rainy image. $\hat{\mathbf{I}}_i$ is the ground-truth rain-free image in D.

Inspired by the work in [27], we design our detail repair network based on the structure detail context aggregation block (SDCAB). The difference from [27] is that we adopt SDCAB into the whole network flow to make the best use of multi-scale features , while [27] only applied multiscale dilated block in the first layer to extract the image features. We have validated that this modification benefits our detail recovery network. Specifically, SDCAB consists of different scales of dilation convolutions and 1×1 convolutions as shown in Fig. 3(d). Since a large receptive field is very helpful to acquire much contextual information [16], we present 3 dilated convolutions whose dilation scales are 1, 3 and 5 in SDCAB. Then, in order to extract the most important features, we concatenate the output of dilated convolutions and utilize the 1×1 convolution to reduce the feature dimensions. For reducing the complexity in training, the residual network is also introduced into SDCAB.

As shown in Fig. 3(d), the dilated convolution concatenation layer (DCCL) can be expressed as

$$\begin{aligned} DCCL = Conv_{1\times1}(Cat[&Conv_{3\times3,d_1}(X), \\ &Conv_{3\times3,d_3}(X), Conv_{3\times3,d_5}(X)]), \end{aligned} \tag{4}$$



(a) Rainy image 5.74/ 0.2440 | (b) Result of DRRN **X** 23.71/ 0.9024 | (c) Final result **Y** 25.79/ 0.9302 | (d) Ground-truth image Inf / 1 | (e) Image **Y-X**

Figure 4. Image deraining results tested in the dataset of Rain200L. From (a)-(e): (a) the input rainy image, (b) the result **X** by only using the rain residual network (i.e., without the detail repair network), (c) the result **Y** by the DRD-Net, (d) the ground-truth image, and (e) the image of **Y-X** (note: we have inverted the image **Y-X** for better visualization).

where $Conv_{x\times x,d_y}$ denotes the dilated convolutions with the kernel size of $x \times x$, and the dilation scale is y. $Cat(\cdot)$ is a concatenating operation and X is the input feature.

Mathematically, SDCAB can be formulated as

$$SDCAB = Add[X_{input}, BN(DCCL_2)], \tag{5}$$

where $DCCL_2$ is described as

$$DCCL_2 = PRelu(BN(DCCL_1(X_{input}))), \tag{6}$$

A large receptive field plays an important role in obtaining more information. With a larger receptive field, we can obtain more context information, which is helpful to find back the lost details. We can observe from Fig. 4 that, DRD-Net has found back the details that were lost by filtering the rainy image to obtain X. We have provided more experimental results on three synthetic datasets to compare the performance of image deraining with and without the additional detail repair network (DRN) in Table 1. As shown in Table 1, our DRD-Net outperforms other network architectures thanks to its capability to nd back the lost details.

3.3. Comprehensive Loss Function

As mentioned above, we employ the simplest L_2 loss as our objective function. The comprehensive loss function of our two-sub networks can be formulated as

$$\begin{aligned} Loss_{total} = \lambda_1 \sum_{i\in N(D)} ||f(\mathbf{O}_i) - \hat{\mathbf{R}}_i||^2 \\ + \lambda_2 \sum_{i\in N(D)} ||(\mathbf{I}_{p,i} + g(\mathbf{O}_i)) - \hat{\mathbf{I}}_i||^2, \end{aligned} \tag{7}$$

where $\mathbf{O}_i$ denotes the i-th input rainy image, $\mathbf{I}_{p,i}$ denotes the preliminarily derained image obtained by subtracting the rain streaks $\mathbf{R}_i$ from $\mathbf{O}_i$, $\hat{\mathbf{R}}_i$ and $\hat{\mathbf{I}}_i$ are the rain streaks image and the rain-free image respectively, λ_1 and λ_2 are two parameters to balance the two sub-loss functions, which in our experiments are fixed to be 0.1 and 1.0 respectively.

4. Experiment and Discussions

Three synthetic datasets and a real-world dataset are used to validate our DRD-Net.

Synthetic Datasets: On account of the difficulty in acquiring the rainy/clean image pair datasets, we use the synthetic datasets to train our network. [31] provides a synthetic dataset named Rain800, which contains 700 training images and 100 testing image. [28] collects and synthesizes 2 datasets, including Rain200L and Rain200H. Both Rain200L and Rain200H consist of 1800 training images and 200 test images.

Real-world Datasets: [25], [28] and [31] supply some real-world rainy images to validate the robustness of deraining methods. We use those images for objective evaluation.

Training Details: We set the total number of epochs to be 120, and each epoch includes 1000 iterations. During training, we set the depth of our network to be 35, and utilize the non-linear activation PRelu [8]. For optimizing our network, the Adam [13] is adopted with a min-batch size of 4 to train the network. We initialize the learning rate as 0.01, which is divided by 2 every 15 epochs. All the experiments are performed by using an Nvidia 2080Ti GPU.

4.1. Ablation Study

Ablation Study on Different Components: To explore the effectiveness of our DRD-Net, it is necessary to decompose its full scheme into different parts and even replace the network architecture for the ablation study.

- **BL:** Baseline (BL) indicates the residual network without the SE operation, which learns a function that maps the rainy images to the rain streaks.
- **BL+SE:** Adding the SE operation to the baseline.
- **BL+SE+DB:** Employing two sub-networks for image deraining. One network is the rain residual net-

Table 1. Quantitative comparison between our DRD-Net and other network architectures.

Dataset	Metrics	BL	BL+SE	BL+SE+DB	BL+SE+RB	DRD-Net (Ours)
Rain200L	PSNR	35.57	36.17	36.89	37.04	**37.15**
	SSIM	0.9759	0.9778	0.9792	0.9860	**0.9873**
Rain200H	PSNR	26.20	26.49	27.16	27.01	**28.16**
	SSIM	0.8245	0.8473	0.9158	0.9061	**0.9201**
Rain800	PSNR	25.83	26.04	26.09	26.12	**26.32**
	SSIM	0.8093	0.8181	0.8903	0.8966	**0.9018**

work (BL+SE), and the another is detail repair network based on the direct block (DB, see in Fig. 3(a)).

- **BL+SE+RB:** DB is replaced with residual block (RB) in the detail repair network.
- **BL+SE+SDCAB:** Our DRD-Net, which comprises the rain residual network (BL+SE) and the detail repair network based on the proposed structure detail context aggregation block (SDCAB).

Table 2. Ablation study on different settings of our method on the synthetic dataset Rain200H. M denotes the number of feature maps in our network and D is the total depth of our network.

	Metrics	M = 16	M = 32	M = 64
D = 8+3	PSNR	26.36	26.77	26.97
	SSIM	0.9085	0.9117	0.9135
D = 12+3	PSNR	26.52	26.89	27.31
	SSIM	0.9092	0.9129	0.9152
D = 16+3	PSNR	26.93	27.61	28.16
	SSIM	0.9127	0.9183	0.9201

Analysis on SE and SDCAB: To validate the necessity of the structure in Fig. 3(d), we remove the SE operation from the network and show the results in Table 1. It is found that the performance of deraining without the SE operation suffers from slight degradations. This certifies the necessity of the SE operation from another side. In order to evaluate the effectiveness of the SDCAB, we compare our network with other connection style blocks, including the direct block (DB), the residual block (RB) which has been used in DDN [5]. For fair comparisons, we replace SDCAB with DB and RB respectively, which is shown in Table 1. The full scheme of BL+SE+SDCAB outperforms other architectures in the three datasets, which certifies that SDCAB is essential to detail-recovery image deraining.

Ablation Study on Parameter Settings: Results under different parameter settings of DRD-Net can be found in Table 2. We have discussed the effects of the number of feature maps and SDCAB or the rain residual blocks (RRB).

4.2. Comparison with the State-of-the-Arts

We compare our method with several state-of-the-art deraining methods, including 2 prior-based methods, i.e., GMM [17], and DSC [20], and 4 learning-based methods, i.e., DDN [5], RESCAN [16], UGSM [3], DAF-Net [10], SPA-Net [25] and PReNet [21]. All these methods are performed in the same training and testing datasets for fair comparisons.

Our DRD-Net can effectively avoid image degradations caused by deraining as demonstrated in Fig. 5 (the upper row). Although most approaches can remove the rain streaks from the rainy image, the halo artifacts and color distortion have appeared after deraining.

Moreover, it is challenging for most approaches to maintain/recover the details from heavy rainy images as shown in Fig. 5 (the middle and bottom rows). The white stripes of zebra and the bicycle are blurred severely by the compared approaches while they are considered as image details and recovered well by our DRD-Net.

In order to validate the practicability of our DRD-Net, we visually evaluate its performance on a series of real-world rainy images in Fig. 6. DRD-Net can effectively remove the real-world rain streaks from the images while preserving their details, but other approaches somewhat tend to over-smooth the images.

Moreover, the visual comparisons are commonly consistent to the numerical evaluations, which are shown in Table 3. Our DRD-Net mostly obtains the higher values of PSNR and SSIM than other methods on those three datasets.

4.3. Detail Recovery for Other Deraining Networks

Existing deep learning-based deraining methods resort to delicate network design to meet the challenging goal of removing rain streaks but retaining details of similar properties. In contrast, our DRD-Net decomposes this conflicting task into **remove** and **repair** by two parallel network branches, which share the same input and collaborate to spit an high-fidelity output. Apparently, the choice of the rain removal part is not unique, the detail recovery branch can be easily attached to existing deraining networks to boost their performance.

Taking DDN [5] as an example, we experiment with the parallel network consisting of DDN and our detail repair network. For fair comparison, we keep most parameters from the original DDN untouched. The depth and the number of feature channels of the detail repair network are set as 24 and 16 respectively. We randomly select 20 image patches with the size of 64×64 to train the network, which

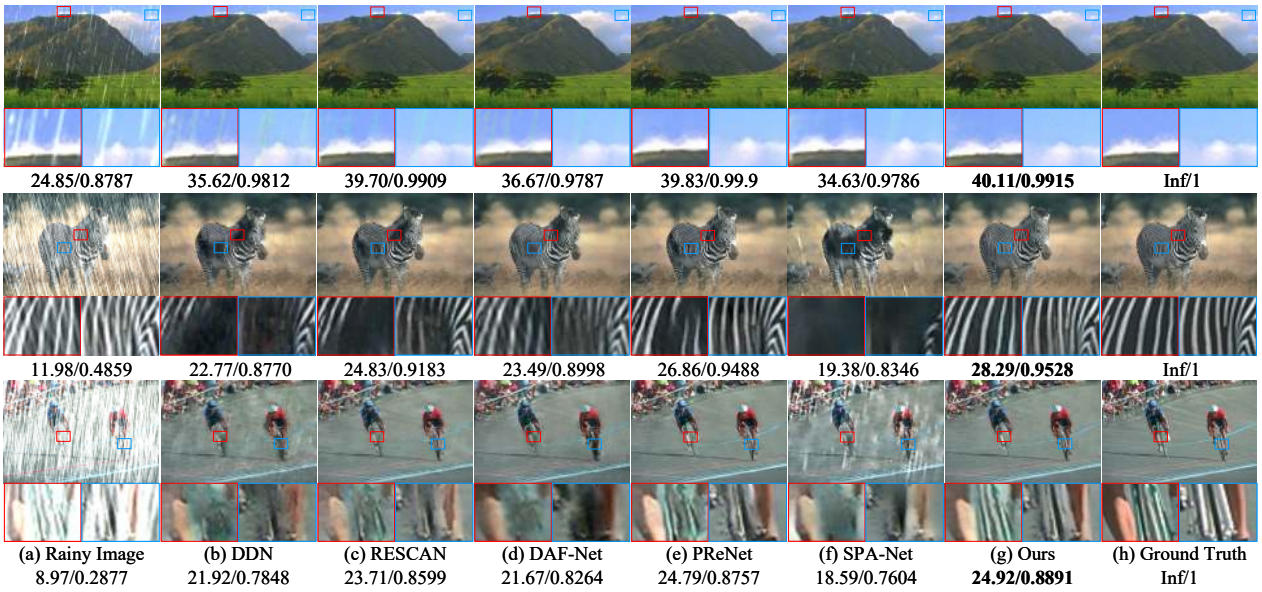


Figure 5. Image deraining results tested in the synthetic datasets. From (a)-(h): (a) the rainy images, and the deraining results of (b) DNN [5], (c) RESCAN [16], (d) DAF-Net [10], (e) PReNet [21], (f) SPA-Net [25], (g) our DRD-Net respectively, and (h) the ground truth.

Figure 6. Image deraining results tested in the real datasets. From (a)-(h): (a) the rainy images, and the deraining results of (b) GMM [17], (c) DDN [5], (d) RESCAN [16], (e) DAF-Net [10], (f) PReNet [21], (g) SPA-Net [25] and (h) our DRD-Net respectively.

is similar to DDN. We evaluate the network in the datasets Rain800 and Rain200H as shown in Table 4. One can observe that DDN incorporated with detail repair network outperforms the original DDN with negligible efficiency sacrifice, thanks to the parallel structure for detail recovery.

Table 4. Quantitative evaluation, DDN w DRN indicates DDN incorporated with the detail repair network.

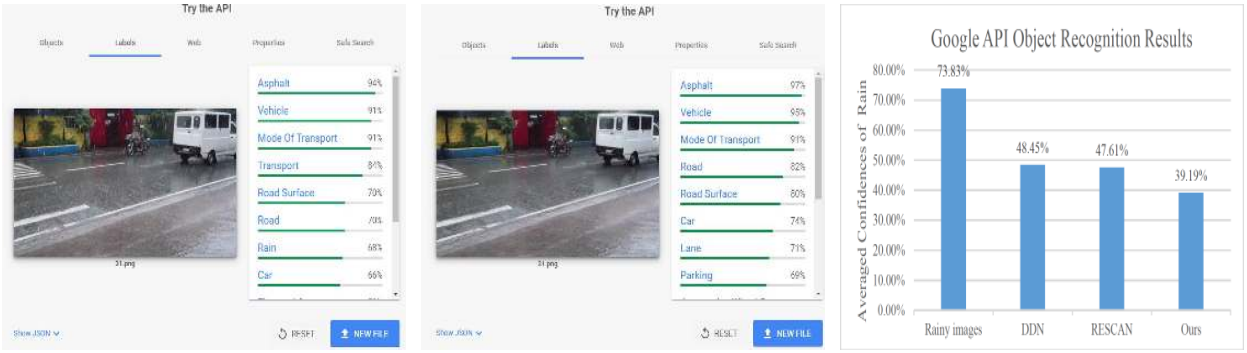
Datasets	Metrics	DDN	DDN w DRN
Rain200H	PSNR	24.64	25.92
	Time	0.03s	0.15s
Rain800	PSNR	24.04	25.13
	Time	0.05s	0.14s

4.4. Running Time

We compare the running time of our method with different approaches on the dataset of Rain200H as shown in Table 5. It is observed that our method is not the fastest one, but its performance is still acceptable.

Table 3. Quantitative experiments evaluated on three recognized synthetic datasets. The first and second best results have been boldfaced and underlined.

Dataset	Rain200L		Rain200H		Rain800	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
GMM [17]	27.16	0.8982	13.04	0.4673	24.04	0.8675
DSC [20]	25.68	0.8751	13.17	0.4272	20.95	0.7530
DDN [5]	33.01	0.9692	24.64	0.8489	24.04	0.8675
RESCAN [16]	<u>37.07</u>	<u>0.9867</u>	26.60	<u>0.8974</u>	24.09	0.8410
DAF-Net [10]	32.07	0.9641	24.65	0.8607	25.27	0.8895
SPA-Net [25]	31.59	0.9652	23.04	0.8522	22.41	0.8382
PReNet [21]	36.76	0.9796	<u>28.08</u>	0.8871	**26.61**	<u>0.9015</u>
Ours	**37.15**	**0.9873**	**28.16**	**0.9201**	<u>26.32</u>	**0.9018**



(a) (b) (c)

Figure 7. The deraining results tested on the Google Vision API. From (a)-(c): (a) object recognition result in the real-world rainy image, (b) object recognition result after deraining by our DRD-Net, and (c) the averaged confidences in recognizing rain from 30 sets of the real-world rainy images and derained images of DDN [5], RESCAN [16] and Our DRD-Net respectively. Note: zero confidence refers to a total failure in recognizing rain from a derained image by Google API.

Table 5. Averaged time (in seconds) and performance of different methods in the dataset of Rain200H.

Metrics	GMM	DSC	DDN	RES CAN	DAF Net	PRe Net	SPA Net	Ours
PSNR	13.04	13.17	24.64	26.60	24.65	28.08	23.04	28.16
Avg time	331.4s	92.9s	0.03s	0.25s	0.52s	0.20s	0.06s	0.54s

4.5. Application

To demonstrate that our DRD-Net can benefit vision-based applications, we employ Google Vision API to evaluate the deraining results. One of the results is shown in Fig. 7 (a-b). It is observed that the Google API can recognize the rainy weather in the rainy image while it cannot recognize the rainy weather in the derained image. Furthermore, we use the Google API to test 30 sets of the real-world rainy images and derained images of our method and two baseline methods [16, 5] as shown in Fig. 7 (c). As one can see, after deraining, the confidences in recognizing rain from the images are significantly reduced.

5. Conclusion

We have presented an end-to-end network with two sub-networks for image deraining from single images. One network is designed to remove the rain streaks from the rainy images, the other is proposed to find back the details to the derained images. We propose the new structure detail context aggregation block (SDCAB) which has a large receptive field to obtain more spatial information. Moreover, qualitative and quantitative experiments indicate that our method outperforms the state-of-the-art learning-based and traditional approaches in terms of removing the rain streaks and recovering the image details.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 61502137, No. 61772267), the HKIBS Research Seed Fund 2019/20 (No. 190-009), the Research Seed Fund (No. 102367) of Lingnan University, Hong Kong, the Fundamental Research Funds for the Central Universities (No. NE2016004), and the Natural Science Foundation of Jiangsu Province (No. BK20190016).

References

[1] Peter C. Barnum, Srinivasa G. Narasimhan, and Takeo Kanade. Analysis of rain and snow in frequency space. *International Journal of Computer Vision*, 86(2-3):256–274, 2010.

[2] Jérémie Bossu, Nicolas Hautière, and Jean-Philippe Tarel. Rain or snow detection in image sequences through use of a histogram of orientation of streaks. *International Journal of Computer Vision*, 93(3):348–367, 2011.

[3] Liang-Jian Deng, Ting-Zhu Huang, Xi-Le Zhao, and Tai-Xiang Jiang. A directional global sparse model for single image rain removal. *Applied Mathematical Modelling*, 59:662–679, 2018.

[4] Zhiwen Fan, Huafeng Wu, Xueyang Fu, Yue Huang, and Xinghao Ding. Residual-guide network for single image deraining. In *2018 ACM Multimedia Conference on Multimedia Conference*, pages 1751–1759. ACM, 2018.

[5] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John W. Paisley. Removing rain from single images via a deep detail network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 1715–1723, 2017.

[6] Kshitiz Garg and Shree K. Nayar. Detection and removal of rain from videos. In *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004), with CD-ROM, 27 June - 2 July 2004, Washington, DC, USA*, pages 528–535, 2004.

[7] Xianglin Guo, Xingyu Xie, Guangcan Liu, Mingqiang Wei, and Jun Wang. Robust low-rank subspace segmentation with finite mixture noise. *Pattern Recognition*, 93:55–67, 2019.

[8] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*, pages 1026–1034, 2015.

[9] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*, pages 7132–7141, 2018.

[10] Xiaowei Hu, Chi-Wing Fu, Lei Zhu, and Pheng-Ann Heng. Depth-attentional features for single-image rain removal. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[11] De-An Huang, Li-Wei Kang, Min-Chun Yang, Chia-Wen Lin, and Yu-Chiang Frank Wang. Context-aware single image rain removal. In *Proceedings of the 2012 IEEE International Conference on Multimedia and Expo, ICME 2012, Melbourne, Australia, July 9-13, 2012*, pages 164–169, 2012.

[12] Li-Wei Kang, Chia-Wen Lin, and Yu-Hsiang Fu. Automatic single-image-based rain streaks removal via image decomposition. *IEEE Transactions on Image Processing*, 21(4):1742–1755, 2011.

[13] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[14] Guanbin Li, Xiang He, Wei Zhang, Huiyou Chang, Le Dong, and Liang Lin. Non-locally enhanced encoder-decoder network for single image de-raining. *arXiv preprint arXiv:1808.01491*, 2018.

[15] Ruoteng Li, Loong-Fah Cheong, and Robby T Tan. Heavy rain image restoration: Integrating physics model and conditional adversarial learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1633–1642, 2019.

[16] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VII*, pages 262–277, 2018.

[17] Yu Li, Robby T. Tan, Xiaojie Guo, Jiangbo Lu, and Michael S. Brown. Rain streak removal using layer priors. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27-30, 2016*, volume 12, pages 2736–2744, 2016.

[18] Xing Liu, Masanori Suganuma, Zhun Sun, and Takayuki Okatani. Dual residual networks leveraging the potential of paired operations for image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7007–7016, 2019.

[19] Yu Luo, Yong Xu, and Hui Ji. Removing rain from a single image via discriminative sparse coding. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*, volume 8, pages 3397–3405, 2015.

[20] Yu Luo, Yong Xu, and Hui Ji. Removing rain from a single image via discriminative sparse coding. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*, pages 3397–3405, 2015.

[21] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: a better and simpler baseline. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3937–3946, 2019.

[22] Varun Santhaseelan and Vijayan K Asari. Utilizing local phase information to remove rain from video. *International Journal of Computer Vision*, 112(1):71–89, 2015.

[23] Abhishek Kumar Tripathi and Sudipta Mukhopadhyay. Removal of rain from videos: a review. *Signal, Image and Video Processing*, 8(8):1421–1430, 2014.

[24] Guoqing Wang, Changming Sun, and Arcot Sowmya. Erl-net: Entangled representation learning for single image deraining. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5644–5652, 2019.

[25] Tianyu Wang, Xin Yang, Ke Xu, Shaozhe Chen, Qiang Zhang, and Rynson WH Lau. Spatial attentive single-image deraining with a high quality real rain dataset. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 12270–12279, 2019.

[26] Wei Wei, Deyu Meng, Qian Zhao, Zongben Xu, and Ying Wu. Semi-supervised transfer learning for image rain re-

moval. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3877–3886, 2019.

[27] Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Joint rain detection and removal via iterative region dependent multi-task learning. *CoRR*, abs/1609.07769, 2016.

[28] Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep joint rain detection and removal from a single image. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 1685–1694, 2017.

[29] He Zhang and Vishal M. Patel. Convolutional sparse and low-rank coding-based rain streak removal. In *2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017, Santa Rosa, CA, USA, March 24-31, 2017*, pages 1259–1267, 2017.

[30] He Zhang and Vishal M Patel. Density-aware single image de-raining using a multi-stream dense network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 695–704, 2018.

[31] He Zhang, Vishwanath Sindagi, and Vishal M. Patel. Image de-raining using a conditional generative adversarial network. *CoRR*, abs/1701.05957, 2017.

[32] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.

[33] Lei Zhu, Chi-Wing Fu, Dani Lischinski, and Pheng-Ann Heng. Joint bi-layer optimization for single-image rain streak removal. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 2545–2553, 2017.