Applied Intelligence
https://doi.org/10.1007/s10489-020-02056-w

Single image rain streak removal via layer similarity prior

Wanshu Fan[1,2] · Yutong Wu[2] · Cong Wang[2]

Accepted: 31 October 2020

Abstract

Single image rain streak removal is a significant and challenging task, which is widely applied in many artificial intelligence domains as preprocessing process. Most of existing rain streak removal works focus on designing various deraining unit (e.g., multi-stream dilation convolution) without considering the correlation between different convolution layers, which may lead to large model size. In this paper, we propose a simple and effective deep network architecture for single image rain streak removal based on deep Convolutional Neural Network (CNN). Benefit from the adjacent layers with different dilation factors have similar feature structures, we design a powerful rain streak representation network based on the Layer Similarity Prior Block (LSPB). To better cater to the property of layer similarity prior, the multi-dense-short-connection is developed and it regards every LSPB as a convolution layer, this connection style makes our overall framework to be a layer similarity prior network. To the best of our knowledge, this is the first paper to investigate the effectiveness of exploiting the layer similarity prior and the multi-dense-short-connection. Quantitative and qualitative experimental results demonstrate that the proposed method outperforms other state-of-the-art methods with the least parameters.

Keywords Deraining · Deep-learning · Layer similarity prior · Multi-dense-short-connection

1 Introduction

Image processing is a hot topic in the field of artificial intelligence, which has been researched everywhere in practical vision tasks. As a classical problem in image processing communities, single image rain streak removal has attracted much attention. Rainy images often degrade the visibility and make the background scene misty, which seriously influence the accuracy of many computer vision systems, especially for some image understanding and recognition tasks, e.g., object detection [19], object tracking [35] and video surveillance. So it is necessary to develop an effective deraining algorithm.

Existing rain streak removal methods can be divided into two categories: video-based methods and single image-based methods. As video-based methods [1, 23, 24, 34] can leverage temporal information by analyzing the difference between adjacent frames, they are easier than single image-based methods. In this paper, we focus on rain streak removal problems for a single image.

Single image rain streak removal aims to recover a clean image from the observed rainy image, usually by estimating the rain streak and then removing them. Mathematically, one rainy image can be modeled as the linear combination of a rain-free image and a rain streak layer, which is usually expressed as

$$O = B + R. \tag{1}$$

where O, B and R denote a rainy image, a clean background image (also called rain-free image) and a rain streak layer, respectively. For single image rain streak removal, recovering B from (1) is a highly ill-posed problem, because many different pairs of B and R give raise to the same O, theoretically.

To make this problem well-posed, most traditional methods attempt to make assumptions on the clean image and the rain streak, such as sparse codeing [21], low-rank representation [3] and Gaussian mixture model [18], may be the most widely used priors in early deraining methods. Although these prior based methods have shown good performance, these priors usually assume that the rain steaks should be sparse and have similar characters in falling

✉ Wanshu Fan
fan921amber@163.com

[1] School of software Engineering, Dalian University, Dalian, China

[2] School of Mathematical Sciences, Dalian University of Technology, Dalian, China

Published online: 16 January 2021

directions and shapes, which do not generalize well for some complex cases and rain forms.

Recently, Convolutional Neural Networks (CNNs) have achieved tremendous progress in many computer vision tasks, e.g., object detection [29], object tracking [35], semantic segmentation [20], super resolution [4, 5] and style transfer [8, 11]. Benefiting from the powerful ability of feature representation of CNNs, image rain streak removal methods based on CNNs have been proposed [6, 7, 15–17, 25–28, 31–33]. These deep-learning based methods perform better than conventional prior based methods by a large margin, but they do not emphasize the inner correlation between different layers, which lead to the large model size. An effective deraining method with small model size, which can also remove various kinds of rain streak, is expected.

Previous methods focus on designing various deep networks (e.g., multi-stream dilation convolution [31], residual network [7]) to learn the transformation from the input image to the rain streak and neglect the inner property of different layers, which will lead to large model size. Motivated by addressing this issue, we explore two important aspects in the design of rain streak removal network: (a) the property of layer similarity, and (b) multi-dense-short-connection. On the one hand, we find that different layers with different dilation factors have similar feature structures, which reflected on paying attention to either clean background or rain streak. Based on the property of layer similarity, we design the Layer Similarity Prior Block (LSPB) to fuse these similar feature structures to maintain the primary information of the rain streak. On the other hand, symmetry skip-connection between corresponding feature maps is widely utilized to transmit low-level features to high-level semantic features, which enable the computation of long-range spatial dependencies as well as efficient usage of the feature activation of proceeding layers. However, we also find that the symmetry skip connection does not fit the layer similarity in a global way. We attempt to develop a more effective connection style: multi-dense-short-connection, it regards every LSPB as a convolution layer, which cater to the property of layer similarity from a global perspective.

Our main contributions are summarized as follows:

- We analyze that there is an inner correlation between different layers with different dilation factors, which makes the different layers have similar feature structures. Based on this property, we propose the LSPB as the basic deraining unit. In addition, this finding also provides some prior knowledge for future deraining algorithm.
- We develop the multi-dense-short-connection to connect several LSPBs, which may cater to the property of layer similarity, and we also verify this connection style is more effective than symmetry skip connection.
- Quantitative and qualitative experimental evaluations on both synthetic datasets and real-world rainy images demonstrate that the proposed network is able to achieve the state-of-the-art results with the least parameters.

2 Related work

In this section, we briefly review several recent related single image rain streak removal methods. In general, single image rain streak removal methods can be grouped into two categories, prior based methods, and deep-learning based methods.

Prior based methods As discussed in Section 1, prior based methods need additional constraints to solve single image rain streak removal problem. Hence, various kinds of prior knowledge are enforced into the optimization framework. In early deraining methods, the most widely used priors are image decomposition [12], non-local filter [13], sparse codes [21], low-rank model [3], and Gaussian mixture model [18]. Kang et al. [12] regard rain streak as high frequency structures and decompose a rain image into the low-frequency and high-frequency layer using the bilateral filter. Then, they decompose high-frequency parts into a rain component and a nonrain component by dictionary learning and sparse coding. Kim et al. [13] detect rain streak regions by analyzing the rotation angle, the aspect ratio of the elliptical kernel at each pixel location, and then perform the nonlocal means filtering on the detected rain streak regions by selecting nonlocal neighbor pixels and their weights adaptively. Luo et al. [21] propose a discriminative sparse coding framework based on image patches and separate rain streak regions from rain-free background images. Chen and Hsu et al. [3] consider the rain streak usually reveal similar and repeated patterns on imaging scene, then they design a low-rank model from matrix to tensor structure for capturing the correlated rain streak and utilize the model to remove rain streak from image in a unified way. Li et al. [18] attempt to use patch-based priors to separate the rain streak from background images. These priors are based on Gaussian mixture models and can accommodate multiple orientations and scales of the rain streak. All the above methods rely on design appropriate priors based on the observations of specific clean images or rain streak properties, they tend to have unsatisfactory performances on real images with complicated scenes and rain forms.

Deep-learning based methods From 2017, deep-learning based methods have been made remarkable progress in image rain streak removal problem. These methods aim to learn a non-linear function, which maps the input rainy image to rain streak. Fu et al. [6, 7] firstly introduce deep-learning methods to solve single image rain streak removal problem. They

decompose rainy images O into low- and high-frequency parts separately, and then map high-frequency parts to rain streak R by a deep residual network. Finally, the estimated rain-free image B is obtained by subtraction operation according to (1). Yang et al. [31] propose a recurrent rain detection and removal network, which jointly detects and removes rain streak iteratively and progressively. Li et al. [15] propose a non-locally enhanced encoder-decoder deraining method, the utilized non-locally enhanced dense blocks are designed not only fully exploit hierarchical features from all the convolutional layers but also well capture the long-distance dependencies and structural information. They achieve satisfactory performance on rain removal in synthetic datasets, but their model size is very large. Li et al. [17] propose a Recurrent Squeeze-and-Excitation Context Aggregation Net (RESCAN) to remove rain streak, where they utilize squeeze-and-excitation [9] to allot each learnable value to rain streak with different sizes. Zhang et al. [32] present a novel density-aware multi-stream densely connected convolutional neural network for joint rain density estimation and deraining, the proposed network is able to automatically determine the rain-density information and then efficiently remove the corresponding rain streak guided by the estimated rain density label. Although this method gets better deraining results in real-world datasets, the success of this algorithm relies on the true density label from the classifier. It may fail to remove rain streak if the false density label classification.

3 Proposed method

In this section, we introduce the details of the proposed deraining method. We first describe the overall architecture of the proposed network in Section 3.1 and then we provide the detailed configuration on the layer similarity prior in Section 3.2. Finally, Section 3.3 presents the loss function which is used in the proposed deraining network.

3.1 Overall network framework

The proposed network aims to learn a non-linear f, which directly describes the mapping relationship between rainy image and rain streak, and then estimate the final rain-free image by subtraction operation via (1). Our overall framework shown in Fig. 1 is composed of several LSPBs connected by multi-dense-short-connection, which is based on the observation that adjacent layers with different dilation factors have similar feature structures. Mathematically, our network can be formulated as follows:

First, we convert the image space into the feature space.

$$F_0 = SE(Conv(O)), \tag{2}$$

where O denotes the rainy image. $Conv$ and SE denote convolution with 3×3 kernel size and squeeze-and-excitation operation, respectively. Here, we utilize squeeze-and-excitation operation to adaptively recalibrate the feature response of each feature map and further obtain the semantic correlation between different channels.

Then, we connect these proposed LSPBs by the multi-dense-short-connection, each LSPB can be regarded as a convolution layer , which satisfies the property of layer similarity from global perspective:

$$F_i = LSPB_i(Conv_{1\times1}(Cat[F_{i-1}, F_{i-3}, \cdots, F_l])), \tag{3}$$

where Cat and $Conv_{1\times1}$ denote concatenation operation at the dimension of the channel and 1×1 convolution respectively. $LSPB_i$ and F_i denote the proposed i-th LSPB and corresponding output, respectively. More details about LSPB will be described in Section 3.2. $l = 0$ if i is odd otherwise $l = 1$, and $i = 1, \cdots, 10$.

The rain streak $\tilde{S}$ can be estimated:

$$\tilde{S} = Conv_{1\times1}(SE(Conv(F_{fusion}))), \tag{4}$$

where $F_{fusion} = Conv_{1\times1}(Cat[F_{10}, F_8, \cdots, F_0])$ is the fusion output of multi LSPBs, it can be regarded as the final estimated rain streak.

Finally, we obtain the final estimated rain-free image $\tilde{B}$ via (1):

$$\tilde{B} = O - \tilde{S}. \tag{5}$$

Locally, each LSPB based on that adjacent layers with different dilation factors have similar feature structures. Globally, the proposed connection style, multi-dense-short-connection, regards each LSPB as a convolution layer. Consequently, the overall network is also a layer similarity prior network.

3.2 Layer similarity prior

Motivated by the adjacent convolution layers have similar structures, we design an effective deraining unit based on this property. We can rely these similar structures and further fuse them for extracting main rain streak. To the best of our knowledge, this is the first attempt to investigate the benefit of leveraging the property of layer similarity between adjacent convolution layers. The proposed LSPB is shown in Fig. 2d, which consists of several dilation convolutions [2], squeeze-and-excitation [9] and one 1×1 convolution. Mathematically, the proposed LSPB can be expressed as:

$$y = LSPB(x_0), \tag{6}$$

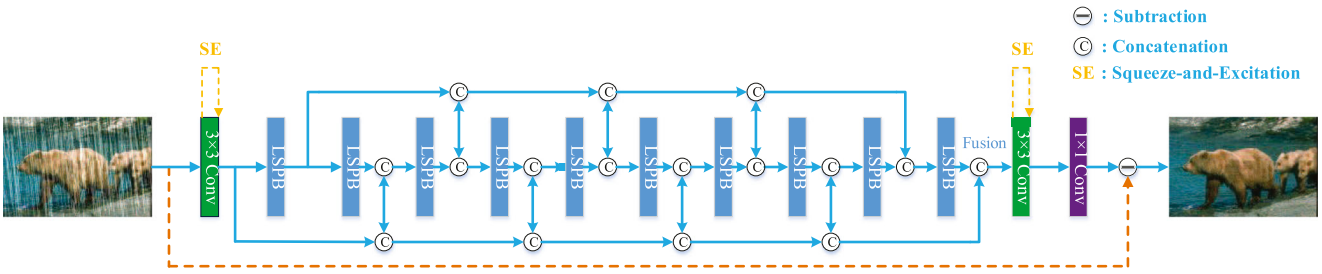

Fig. 1 Layer similarity prior network (**LSPN**). LSPB denotes layer similarity prior block, shown in Fig. 2 (d)

where x_0 denotes the input signal of LSPB, y denotes the output of LSPB, $LSPB$ denotes the operational process of the layer similarity prior.

In the following, we describe the $LSPB$ in details:

First, we utilize several dilation convolutions to obtain more spatial contextual information that has a mass of similar structures between them:

$$x_r = \begin{cases} Conv_r(x_0) & r = 1 \\ Conv_r(x_1) & r = 3, 5 \end{cases}, \tag{7}$$

where $Conv_r$ and x_r denote convolution with dilation factor r and corresponding output, respectively.

Then, we use 1×1 convolution to fuse them for extracting the main rain streak:

$$z_r = \begin{cases} Conv_r(x_1) & r = 1 \\ Conv_{1\times1}(Cat([x_1, x_r])) & r = 3, 5 \end{cases}. \tag{8}$$

Finally, we cascade all layers to make the output pay more attention to rain streak rather than background. The final output of LSPB is

$$y = SE(Conv_{1\times1}(Cat[z_1, z_3, z_5])). \tag{9}$$

A series of operations make the block neglect the background information and pay attention to the rain streak, as shown in Figs. 11 and 12, which makes our network has stronger rain streak representation ability. The analysis on the layer similarity will be discussed in details in Section 4.5. The effect on the model size and the number of LSPBs will be discussed in Section 4.8.

3.3 Loss function

We use MSE as the error function:

$$\mathcal{L} = \frac{1}{HWC} \sum_{t=1}^{H} \sum_{s=1}^{W} \sum_{k=1}^{C} \|\widehat{\boldsymbol{B}}_{t,s,k} - \boldsymbol{B}_{t,s,k}\|_2^2, \tag{10}$$

where H, W, C denote the height, width and channel number of the rain-free image, respectively. $\widehat{\boldsymbol{B}}$ and $\boldsymbol{B}$ denote the estimated rain-free image and the clean background image, respectively. Actually, this loss is equivalent to $\frac{1}{HWC} \sum_{t=1}^{H} \sum_{s=1}^{W} \sum_{k=1}^{C} \|\widehat{\boldsymbol{R}}_{t,s,k} - \boldsymbol{R}_{t,s,k}\|_2^2$ according to the rainy image decomposition (1). $\widehat{\boldsymbol{R}}$ and $\boldsymbol{R}$ denote the estimated rain streak and corresponding ground-truth, respectively.

4 Experimental results

In this section, we demonstrate the effectiveness of the proposed method by conducting various experiments on two synthetic datasets and several real-world images. The experimental settings and criterions of quality evaluation

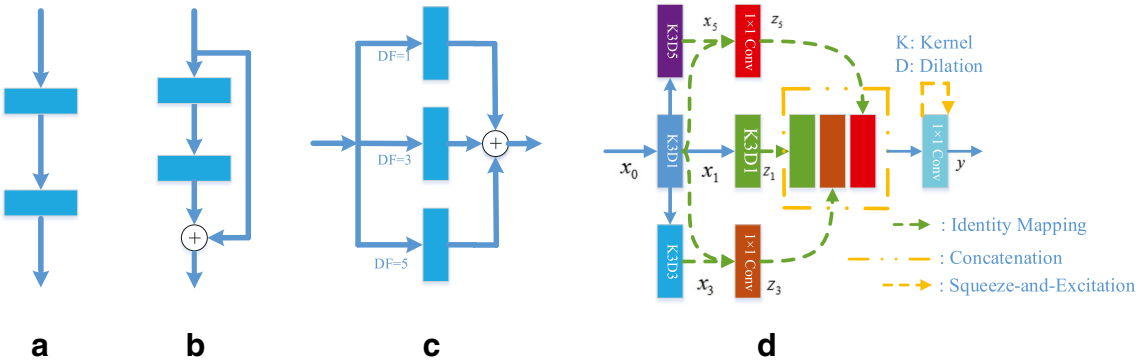

Fig. 2 Some classical convolution styles. **a** Traditional convolution used in DCGAN [33]. **b** Residual block used in DDN [7]. **c** Dilation convolution used in JORDER [31]. **d** Our proposed LSPB

are shown in Section 4.1. Quantitative and visual comparisons on synthetic datasets are reported in Section 4.2. Quantitative and visual comparisons on the real-world rainy images are provided in Section 4.3. An ablation study is conducted in Section 4.4 to demonstrate the effectiveness of the proposed network. Some analysis of the proposed network are presented in Sections 4.5, 4.6, 4.7, and 4.8. The results are compared with five state-of-the-art methods: DSC [21] (ICCV15), LP [18] (CVPR16), DDN [7] (CVPR17), JORDER [31] (CVPR17), and DID [32] (CVPR18).

4.1 Experiment settings

Synthetic datasets We evaluate the performance of the proposed method on two widely used synthetic datasets: Rain100L [31] and Rain100H [31]. Rain100L and Rain100H contain 1800 images for training and 200 images for testing, respectively. These two datasets include various rain streak with different sizes, shapes and directions. Using these synthetic images for training can boost the capacity of the network. We select Rain100H as our analysis dataset in Sections 4.4–4.8.

Real-world datasets Part of the real-world rainy images we use for testing are provided by the previous methods [31, 33]. The other challenging ones are downloaded from the Internet. We randomly pick 50 rainy images from them as our real-world dataset. For the real-world dataset, the rain streak is different from orientation to density. We use them to illustrate the effectiveness of our network.

Measurements Peak signal to noise ratio (PSNR) [10] and structure similarity index (SSIM) [30] are widely used to evaluate the quality of restored results with ground-truth. PSNR is based on the error between corresponding pixels, i.e., estimated deraining result and ground truth. The higher its value, the better the restored image will be. SSIM is a measure of similarity between two images. The closer to 1, the more similar the two images are. We use them as our measurement criteria on synthetic datasets. For the real-world images, we use the perceptual metric Natural Image Quality Evaluator (NIQE) and visual comparisons to evaluate the performance of our method. NIQE is demonstrated to be highly correlated with human ratings and easy to implement. it calculates the no ground-truth image quality score for the real-world rainy image, which are detailed in [22]. A lower NIQE value indicates the better perceptual quality.

Training details We set the number of channels $L = 10$ and the number of LSPBs $M = 10$. We use the LeakyReLU with $\alpha = 0.2$ as the non-linear activation for all convolutions except the last layer. We randomly crop 100×100 patch pairs from training datasets as inputs with a mini-batch size of 10 to train our network. The ADAM [14] is used as the optimization algorithm with an initialized learning rate of 0.001, and the rate will be divided by 10 at $240K$ and $320K$ iterations, and total training iterations are $400K$. We use PyTorch to perform all experiments on an NVIDIA GTX 1080Ti GPU, and the network is trained end-to-end.

4.2 Results on synthetic datasets

Quantitative comparisons between the proposed method and five state-of-the-art deraining methods are shown in Table 1. These deraining methods include two prior based methods: DSC [21] and LP [18], and three deep-learning based methods: DDN [7], JORDER [31], and DID [32]. On the datesets of Rain100H and Rain100L, it can be clearly observed that our method significantly outperforms the two prior based methods in terms of both PSNR and SSIM. Compared with three deep-learning based methods, our method also achieves the highest scores of the evaluative criteria on the two datasets while using the least number of parameters. It's worth mentioning that although our method reduces 87% number of parameters compared with JORDER [31] and DID [32], both of our PSNR and SSIM are superior to these methods.

We provide several examples to compare the proposed method with five state-of-the-art methods visually. Figure 3 shows two synthetic examples compared with two prior based methods, we can see that the prior based methods DSC [21] and LP [18] fail to remove the rain streak and still maintain a significant amount of rain streak, while our method is able to effectively remove rain streak and generates the clearest and cleanest deraining images. Moreover, we also compare the proposed method with three deep-learning based methods DDN [7], JORDER [31], and DID [32] in Fig. 4. DDN [7] and JORDER [31] generate results with some remaining rain streak or artifacts, as shown in Fig. 4b and c, respectively. The result by ID [32] is visually a little better, but it still has some remaining rain streak as shown in Fig. 4d. Compared with it, our method generates the cleanest result with the least artifacts as shown in Fig. 4e.

4.3 Results on real-world images

To further demonstrate the effectiveness of the proposed method, we also evaluate the performance of the proposed method on several real-world rainy images compared with five state-of-the-art methods. Quantitative comparisons between the proposed method and five state-of-the-art deraining methods are shown in Table 2, we find that our method achieves better NIQE values than the other methods

Table 1 Quantitative experiments evaluated on synthetic datasets

	DSC [21]		LP [18]		DDN [7]		JORDER [31]		DID [32]		Ours	
Dataset	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Rain100H	15.66	0.42	14.26	0.54	22.26	0.69	23.45	0.74	<u>26.12</u>	<u>0.83</u>	**26.87**	**0.85**
Rain100L	24.16	0.87	29.11	0.88	34.85	0.95	36.11	<u>0.97</u>	36.14	0.96	**37.18**	**0.98**
Parameters	–		–		58,175(−17 %)		369,792(−87 %)		372,839(−87 %)		**48,185**	

The best and the second best results are boldfaced and underlined, respectively

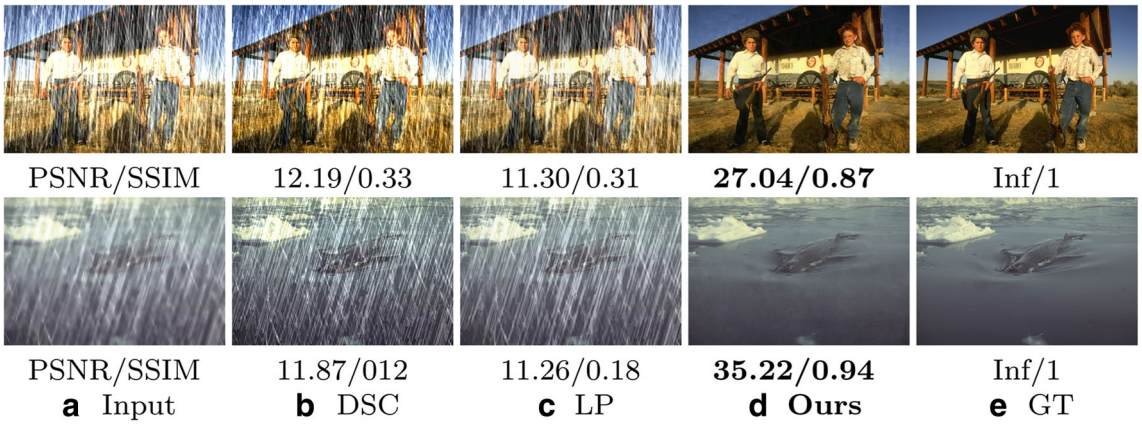

Fig. 3 Two example results on the synthetic datasets compared with two prior based deraining methods. The input image shown in (**a**) is synthesized by large-scale dense rain streak. Our results shown in (**d**) have the highest PSNR and SSIM values. Visually, our method generates the cleanest results

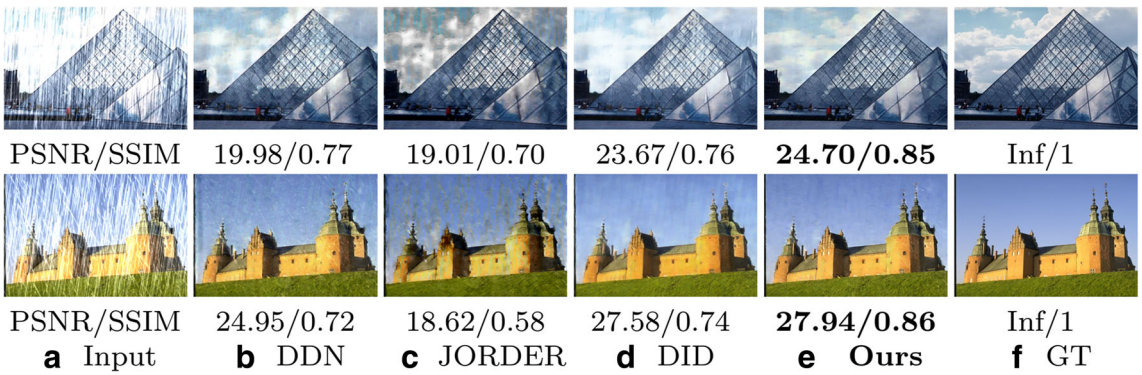

Fig. 4 Two examples on the synthetic datasets compared with three deep-learning based methods. Our results are the best both quantitatively and visually

Table 2 Quantitative experiments evaluated on the real-world dataset

Metric	Input	DSC [21]	LP [18]	DDN [7]	JORDER [31]	DID [32]	Ours
NIQE	5.09	4.89	4.76	4.13	<u>3.77</u>	4.26	**3.59**

The best and the second best results are boldfaced and underlined, respectively

on the real-world images. The results of various deraining methods are shown in Figs. 5 and 6. The numbers in red presented on the top left corner in each image of Figs. 5 and 6 are the NIQEs values, which quantitatively measure the quality of the deraining methods. Figure 5, presents two real-world examples compared with two prior based methods: DSC [21] and LP [18]. It is obvious that the prior based methods tend to have some residual rain streak as shown in Fig. 5b and c. Both quantitatively and visually, our method obtains better NIQE scores and achieves favorable results against the other prior based methods as shown in Fig. 5d. In addition, we test our method on several real-world rainy images from different scenes. To provide visual comparisons, Fig. 6 shows four examples compared with three deep-learning based methods. For the first two examples, our method shown in Fig. 6e generates the cleanest results with clearer background details, while other methods either still tend to leave some rain streak or produce unclear texture. See the close-ups for details. For the last two examples, DDN [7], JORDER [31], and DID [32] can remove the majority of rain streak, but they still remain some rain streak as shown in Fig. 6b, c and d, respectively. In comparison, our method generates the cleanest results with the least artifacts. We also find that our method achieves the best NIQE performance on all test real-world rainy images.

More deraining results by the proposed method are presented in Fig. 7.

4.4 Ablation study

As our network consists of several LSPBs and multi-dense-short-connection, and the LSPB is based on dilation convolution and squeeze-and-excitation operations, it is meaningful to explore the effectiveness of the two operations and the connection style. To this end, we conduct two detailed ablation studies on different methods. The first ablation study is conducted to demonstrate the improvements obtained due to different components in the proposed LSPB. We perform the following experiments on Rain100H with the same settings.

- **BL**: A single baseline network without dilation convolution and squeeze-and-excitation operations.
- **BL + DC**: Baseline network with dilation convolution.
- **BL + SE**: Baseline network with squeeze-and-excitation operation.
- **BL + DC + SE**: Baseline network with dilation convolution and squeeze-and-excitation operations, i.e., the proposed network.

Quantitative evaluations are reported in Table 3. It can be observed that BL performs much worse than the proposed network in terms of PSNR and SSIM (25.91 dB v.s. 26.87 dB, 0.82 v.s. 0.85). By plugging the dilation convolution and squeeze-and-excitation into the baseline network, we can improve the value of PSNR and SSIM by 0.96 dB and 3 %, respectively. Overall, these comparisons firmly indicate the dilation convolution and squeeze-and-excitation operations benefits for performance improvement.

Moreover, we also provide one example for visual comparisons to verify the effectiveness of the dilation convolution and squeeze-and-excitation operations in Fig. 8. It can observe that the result of the final network (**BL+DC+SE**) shown in Fig. 8e is the best both quantitatively and visually.

In the second ablation study, we explore the effectiveness of different connection styles, including no-connection (NC), symmetric skip connection (SSC) and our proposed multi-dense-short-connection (MDSC). The average PSNR and SSIM results evaluated on Rain100H for various connection styles are tabulated in Table 4. We can see that our proposed MDSC is more effective than the other two styles. We also provide two examples for visual comparisons.

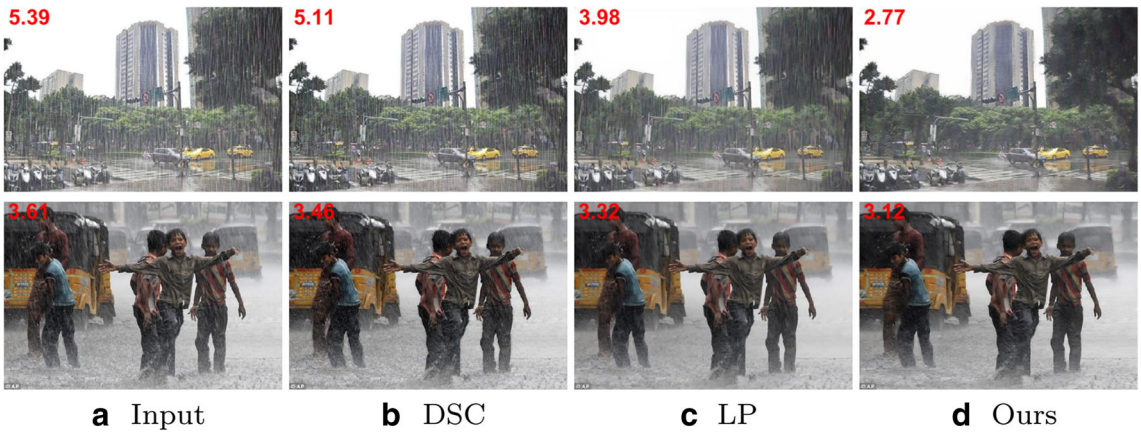

Fig. 5 Two real-world examples compared with two prior based methods. The results of DSC [21] and LP [18] still remain some rain streak. In comparison, our method generates much clearer results and our results have the lowest NIQEs values

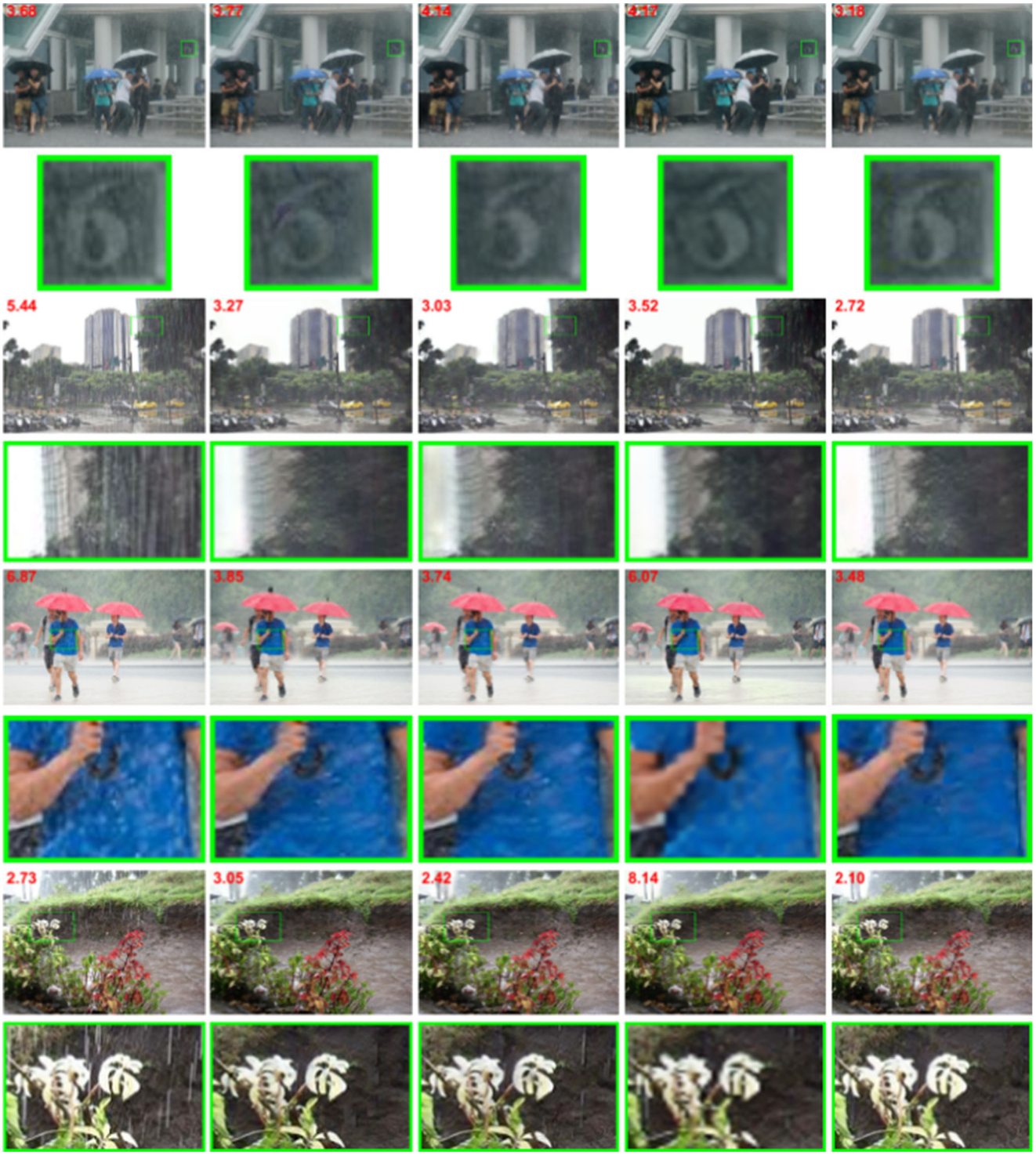

Fig. 6 Several example results with close-ups on the real-world rainy images compared with three deep-learning based methods. Our results have the lowest NIQEs values. Visually, our method generates the cleanest results with the least artifacts

As shown in Fig. 9, the method using no-connection is unable to remove rain streak. Even though the method using symmetric skip connection is able to successfully remove the majority of rain streak, it still tends to leave some rain streak in the derained images. In contrast, our method using the proposed multi-dense-short-connection is capable

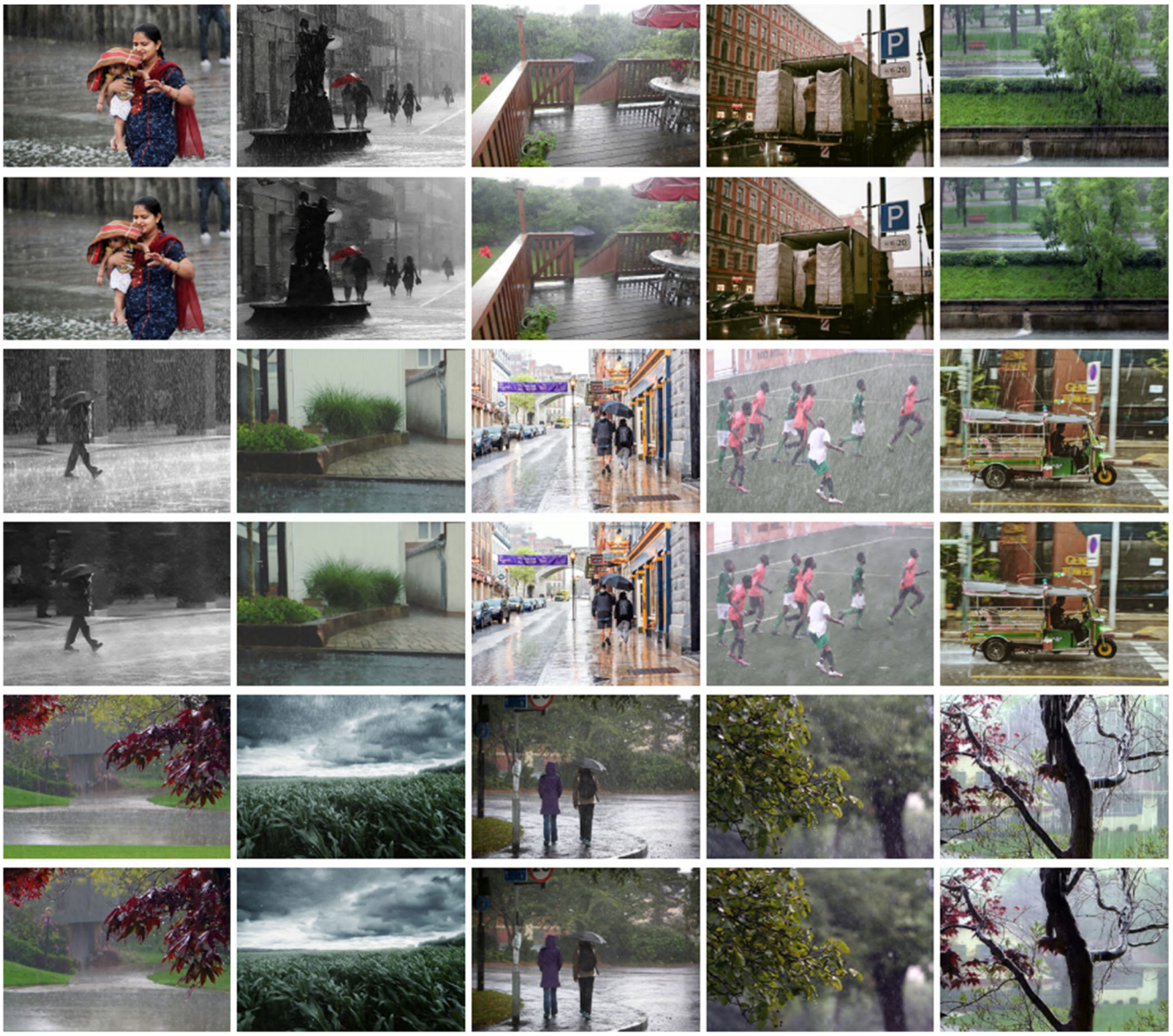

Fig. 7 More results produced by our method on the real-world rainy images under different scenes

of removing rain streak while generating much clearer and cleaner results. Similar observations can be made utilizing the quantitative criterions, PSNR and SSIM, as shown in Fig. 9.

4.5 Analysis on the LSPB

As our network is mainly composed of several LSPBs, a natural question is that whether the LSPB works better than other convolution units used in reported works, e.g., traditional convolution (TC) (basic unit in [33]), residual block (RB) (basic unit in [7]), multi-stream dilation convolution (DC) (used in [31]). To verify the effectiveness of our proposed LSRB, we carry out some experiments by replacing the LSRB with other convolution units. For fair comparison, it is ensured that the number of parameters is almost for all conditions. The results are shown in Table 5. It can be observed that the proposed LSPB obtains the best results while using the least number of parameters among all the compared methods.

Table 3 Ablation study on layer similarity prior block

	BL	$BL+DC$	$BL+SE$	$BL+DC+SE$ (Default)
PSNR	25.91	26.76	26.41	**26.87**
SSIM	0.82	0.84	0.83	**0.85**
Parameters	47,813	47,813	48,185	48,185

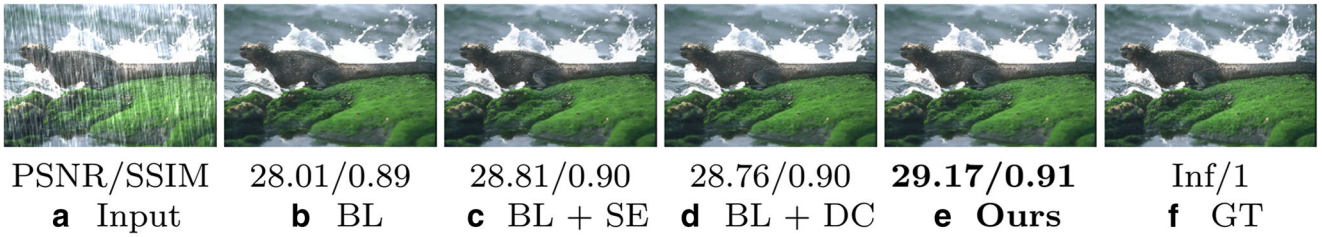

Fig. 8 One visual example for ablation study on the LSPB. Compared with other results, our deraining results shown in (e) has the highest PSNR and SSIM values

Table 4 Ablation study on connection style

	NC	*SSC*	*MDSC*
PSNR	15.56	26.76	**26.87**
SSIM	0.33	0.83	**0.85**
Parameters	45,295	46,345	48,185

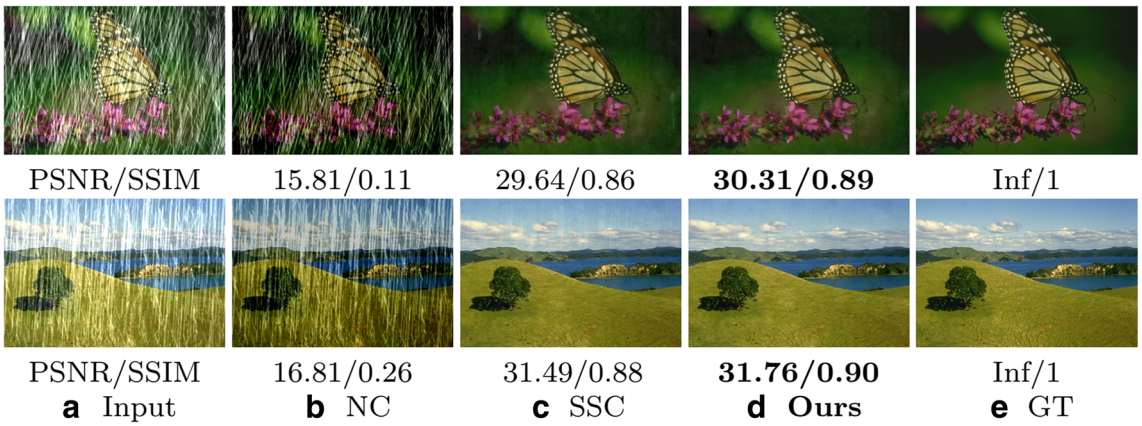

Fig. 9 Visual examples on different connection styles

Table 5 Quantitative experiments evaluated on our proposed LSPB compared with different convolution units

	TC	*RB*	*DC*	LSPB (Ours)
PSNR	26.25	26.20	26.43	**26.87**
SSIM	0.83	0.83	0.83	**0.85**
Parameters	49,835	49,835	52,889	**48,185**

The best result is boldfaced

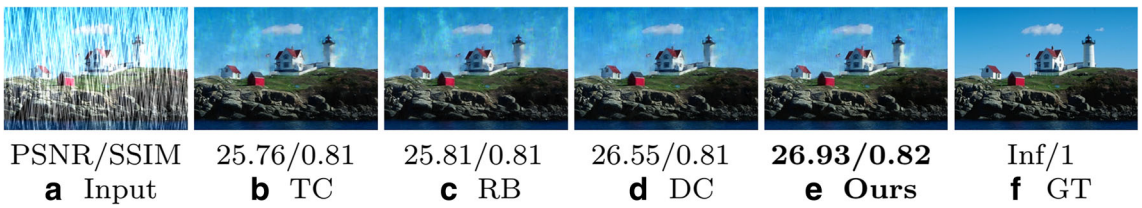

Fig. 10 A visual example of comparisons with different deraining units. Our result shown in (**e**) is the best with least artifacts

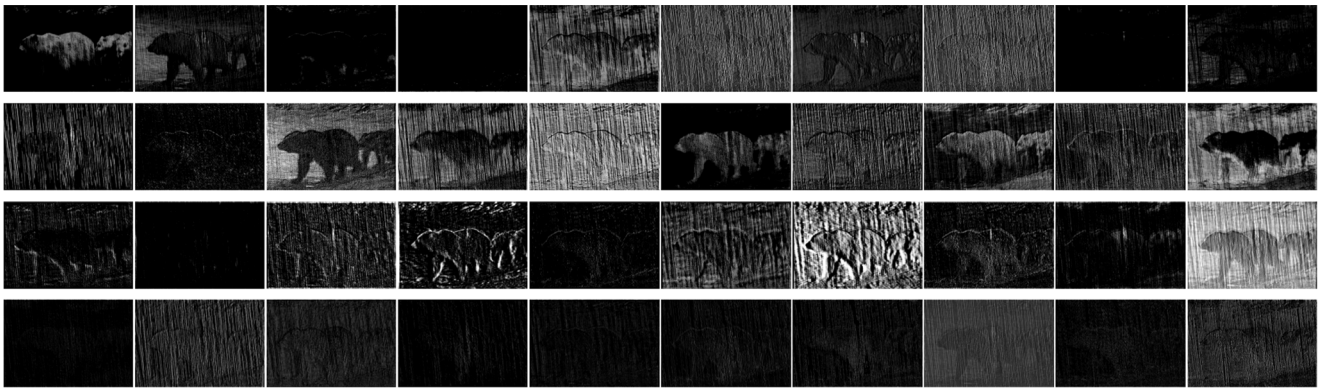

Fig. 11 The visualization of the feature maps in the first LSRB. From top to bottom, $K3D1$, $K3D3$, $K3D5$ and final layer. Please see in high resolution screen!

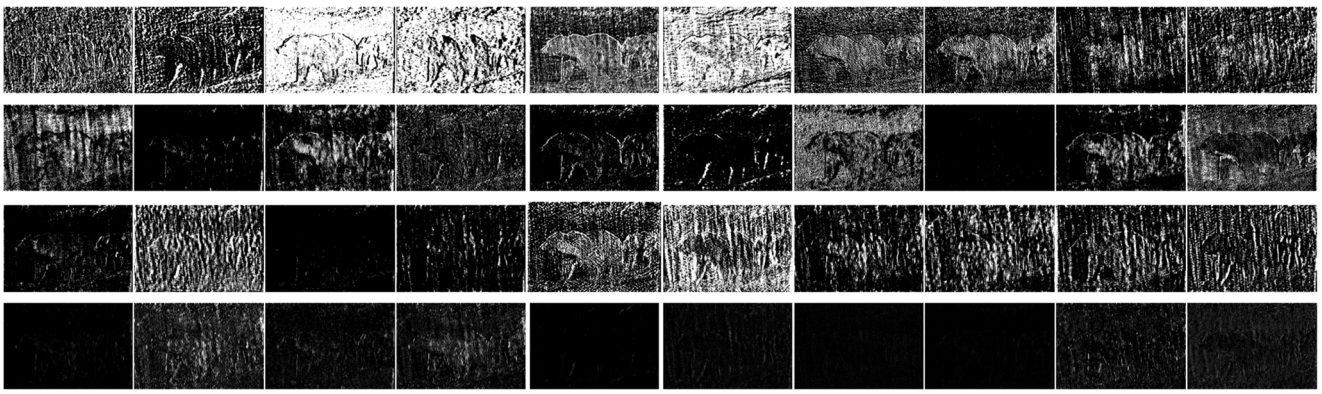

Fig. 12 The visualization of the feature maps in the last LSRB

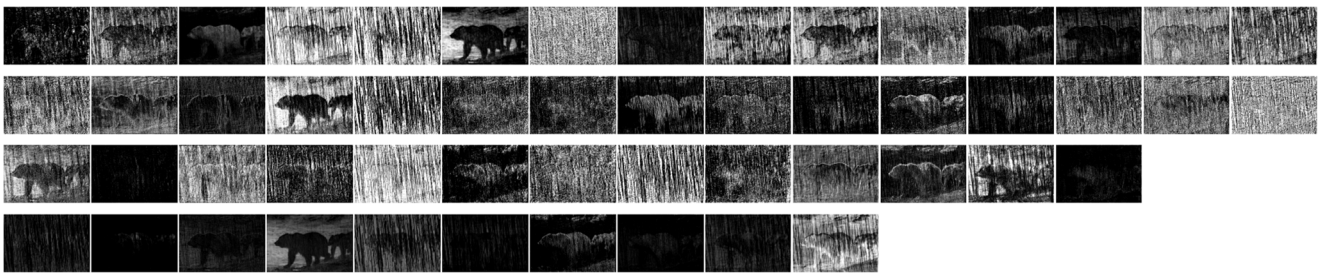

Fig. 13 The visualization of feature maps on the fusion layer. From top to bottom: traditional convolution, residual block, dilation convolution and our deraining unit

Table 6 Quantitative experiments for analyzing the number of channels and LSPBs

	Metrics	$M = 8$	$M = 10$ (Default)	$M = 12$
$L = 8$	PSNR	26.01	26.74	26.99
	SSIM	0.82	0.84	0.85
	Parameters	24,885	38,563	55,225
$L = 10$ (Default)	PSNR	26.54	26.87	27.29
	SSIM	0.83	0.85	0.86
	Parameters	31,079	48,185	69,027
$L = 12$	PSNR	26.71	27.24	27.69
	SSIM	0.84	0.85	0.87
	Parameters	37,337	57,907	82,937

M and L denote the number of channels and LSPBs, respectively

Moreover, an example for visual comparisons with different deraining units is presented in Fig. 10. As can be seen, our proposed deraining unit shown in Fig. 10e achieves the best performance compared with other deraining units. This also illustrates that our proposed LSPB is more effective than other commonly used deraining units.

4.6 The visualization of the LSPB

To better demonstrate the effectiveness of the proposed method layer similarity prior, we present the visualization of the feature maps for the adjacent convolution layers in Figs. 11 and 12. As can be seen from the first three rows of Figs. 11 and 12, the previous layers have similar structures, which take into account either background or rain streak. The last row in Figs. 11 and 12 show the fusion results that only focus on the rain streak rather than the background. This design enables our network to have strong rain streak representation ability.

4.7 The visualization of the fusion layer among different deraining units

We also provide the visualization results on the fusion layer among different deraining units in Fig. 13. It can be observed that our designed LSPB has better expression capacity of rain streak while the other deraining units may fail to accurately capture the rain streak. It is also verify that the proposed deraining unit LSPB is effective for single image rain streak removal. For fair comparison, we adjust the number of channels under the condition of different deraining units, making different deraining units have different number of channels.

4.8 Analysis on the model size and the number of LSPBs

We further analyze the effect of the number of channels and LSPBs. In order to choose suitable numbers of channels and LSPBs, we carry out experiments on rain100H to test the different number of channels and LSPBs. Quantitative results corresponding to the different number of channels and LSPBs are tabulated in Table 6. We can observe that the deraining results become better with the number of channels and LSPB increasing. Combining with the evaluations against the state-of-the-art methods reported in Table 1, we can see that our proposed method is superior to the state-of-the-art methods when we select $L = 12$ and $M = 12$, but the model size is larger. Moreover, the results by our method are also comparable with other state-of-the-art methods when we select $L = 8$ and $M = 8$, and the parameters only have 24,885. As the performance of our method is satisfactory and the number of parameters is less than other state-of-the-art methods when $L = 10$ and $M = 10$, we argue that this setting is good enough. So we choose $L = 10$ and $M = 10$ as our network setting. We also provide an example in Fig. 14 for visual comparisons with

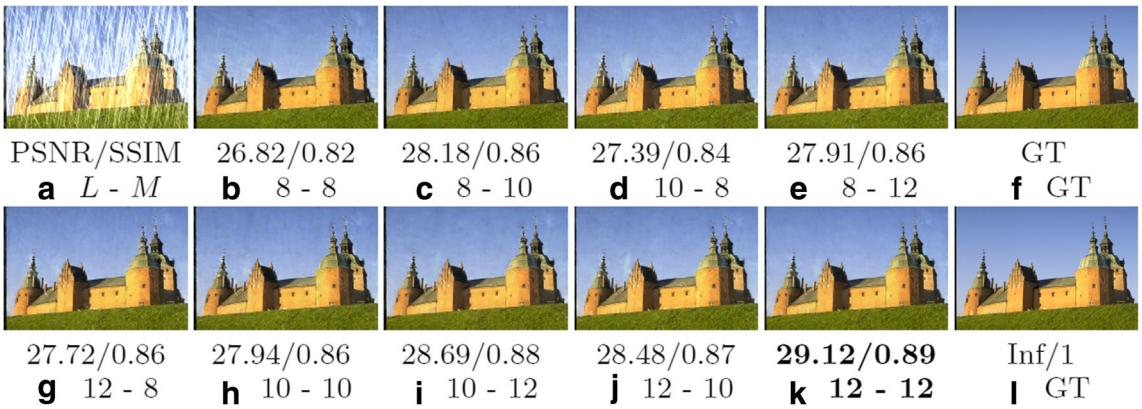

Fig. 14 A visual example of comparisons with different number of channels and LSPBs. M and L denote the number of channels and LSPBs, respectively

different number of channels and LSPBs. As can be seen, the deraining performance gets better with the model size increasing that also illustrates our model is robust.

5 Conclusion

In this paper, we consider the property that adjacent layers with different dilation factors have similar feature structures. Extensive experiments have verified the effectiveness of the proposed layer similarity prior. We utilize the layer similarity prior to design a novel network for single image deraining. The proposed network is constructed based on the LSPB, which can fuse these similar feature structures for better maintain the primary information of the rain streak According to the rainy image decomposition model, the rain-free image can be computed by subtraction. Both quantitative and qualitative experiments evaluated on synthetic and real-world rainy images demonstrate that the proposed method outperforms many recent state-of-the-art deraining methods while using fewer parameters. Moreover, the proposed layer similarity prior and multi-dense-short-connection may also provide prior knowledge for future deraining research and other low-level vision tasks.

Acknowledgments This work was supported by the National Natural Science Foundation of China [grant numbers 61976041]; Major National Science and Technology Project of China [grant number 2018ZX04041001].

References

1. Brewer N, Liu N (2008) Using the shape characteristics of rain to identify and remove rain from video. In: Structural, syntactic, and statistical pattern recognition, pp 451–458, https://doi.org/10.1007/978-3-540-89689-0_49
2. Chen L, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2018) Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. TPAMI 40(4):834–848. https://doi.org/10.1109/TPAMI.2017.2699184
3. Chen Y, Hsu C (2013) A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In: ICCV, pp 1968–1975, https://doi.org/10.1109/ICCV.2013.247
4. Cui Z, Chang H, Shan S, Zhong B, Chen X (2014) Deep network cascade for image super-resolution. In: ECCV, pp 49–64, https://doi.org/10.1007/978-3-319-10602-1_4
5. Dong C, Loy CC, He K, Tang X (2016) Image super-resolution using deep convolutional networks. 38(2):295–307. https://doi.org/10.1109/TPAMI.2015.2439281
6. Fu X, Huang J, Ding X, Liao Y, Paisley J (2017) Clearing the skies: a deep network architecture for single-image rain removal. 26(6):2944–2956. https://doi.org/10.1109/TIP.2017.2691802
7. Fu X, Huang J, Zeng D, Huang Y, Ding X, Paisley J (2017) Removing rain from single images via a deep detail network. In: CVPR, pp 1715–1723, https://doi.org/10.1109/CVPR.2017.186
8. Gonzalez-Garcia A, van de Weijer J, Bengio Y (2018) Image-to-image translation for cross-domain disentanglement. In: NIPS, pp 1294–1305
9. Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: CVPR, pp 7132–7141, https://doi.org/10.1109/CVPR.2018.00745
10. Huynh-Thu Q, Ghanbari M (2008) Scope of validity of psnr in image/video quality assessment. Electron Lett 44(13):800–801. https://doi.org/10.1049/el:20080522
11. Johnson J, Alahi A, Fei-Fei L (2016) Perceptual losses for real-time style transfer and super-resolution. In: ECCV, pp 694–711, https://doi.org/10.1007/978-3-319-46475-6_43
12. Kang L, Lin C, Fu Y (2012) Automatic single-image-based rain streaks removal via image decomposition. 21(4):1742–1755. https://doi.org/10.1109/TIP.2011.2179057
13. Kim J, Lee C, Sim J, Kim C (2013) Single-image deraining using an adaptive nonlocal means filter. In: ICIP, pp 914–917, https://doi.org/10.1109/ICIP.2013.6738189
14. Kingma DP, Ba J (2015) Adam: a method for stochastic optimization. In: ICLR
15. Li G, He X, Zhang W, Chang H, Dong L, Lin L (2018) Non-locally enhanced encoder-decoder network for single image deraining. In: ACM MM, pp 1056–1064, https://doi.org/10.1145/3240508.3240636
16. Li S, Araujo IB, Ren W, Wang Z, Tokuda EK, Junior RH, Cesar-Junior R, Zhang J, Guo X, Cao X (2019) Single image deraining: a comprehensive benchmark analysis. In: CVPR. computer vision foundation / IEEE, pp 3838–3847, https://doi.org/10.1109/CVPR.2019.00396
17. Li X, Wu J, Lin Z, Liu H, Zha H (2018) Recurrent squeeze-and-excitation context aggregation net for single image deraining. In: ECCV, pp 262–277, https://doi.org/10.1007/978-3-030-01234-2_16
18. Li Y, Tan RT, Guo X, Lu J, Brown MS (2016) Rain streak removal using layer priors. In: CVPR, pp 2736–2744, https://doi.org/10.1109/CVPR.2016.299
19. Lin T, Dollár P, Girshick RB, He K, Hariharan B, Belongie SJ (2017) Feature pyramid networks for object detection. In: CVPR, pp 936–944, https://doi.org/10.1109/CVPR.2017.106
20. Long J, Shelhamer E, Darrell T (2015) Fully convolutional networks for semantic segmentation. In: CVPR, pp 3431–3440, https://doi.org/10.1109/CVPR.2015.7298965
21. Luo Y, Xu Y, Ji H (2015) Removing rain from a single image via discriminative sparse coding. In: ICCV, pp 3397–3405, https://doi.org/10.1109/ICCV.2015.388
22. Mittal A, Soundararajan R, Bovik AC (2013) Making a "completely blind" image quality analyzer. IEEE Signal Process Lett 20(3):209–212. https://doi.org/10.1109/LSP.2012.2227726
23. Santhaseelan V, Asari VK (2015) Utilizing local phase information to remove rain from video. IJCV 112(1):71–89. https://doi.org/10.1007/s11263-014-0759-8
24. Tripathi AK, Mukhopadhyay S (2014) Removal of rain from videos: a review. SIViP 8(8):1421–1430. https://doi.org/10.1007/s11760-012-0373-6
25. Wang C, Wang H, Su Z, Yang Y (2019) Embedding non-local mean in squeeze-and-excitation network for single image deraining. In: ICMEW, pp 264–269, https://doi.org/10.1109/ICMEW.2019.00-76
26. Wang C, Zhang M, Pan J, Su Z (2019) Single image rain removal via densely connected contextual and semantic correlation net. JEI 28(3):033,018. https://doi.org/10.1117/1.JEI.28.3.033018
27. Wang C, Zhang M, Su Z, Wu Y, Yao G, Wang H (2019) Learning a multi-level guided residual network for single image deraining. Signal Process Image Commun 78:206–215. https://doi.org/10.1016/j.image.2019.07.003

28. Wang C, Zhang M, Su Z, Yao G, Wang Y, Sun X, Luo X (2019) From coarse to fine: a stage-wise deraining net. IEEE Access 7:84,420–84,428. https://doi.org/10.1109/ACCESS.2019.2922549
29. Wang X, Shrivastava A, Gupta A (2017) A-fast-rcnn: hard positive generation via adversary for object detection. In: CVPR, pp 3039–3048, https://doi.org/10.1109/CVPR.2017.324
30. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. TIP 13(4):600–612. https://doi.org/10.1109/TIP.2003.819861
31. Yang W, Tan RT, Feng J, Liu J, Guo Z, Yan S (2017) Deep joint rain detection and removal from a single image. In: CVPR, pp 1685–1694, https://doi.org/10.1109/CVPR.2017.183
32. Zhang H, Patel VM (2018) Density-aware single image de-raining using a multi-stream dense network. In: CVPR, pp 695–704, https://doi.org/10.1109/CVPR.2018.00079
33. Zhang H, Sindagi V, Patel VM (2019) Image de-raining using a conditional generative adversarial network. TCSVT. https://doi.org/10.1109/TCSVT.2019.2920407
34. Zhang X, Li H, Qi Y, Leow WK, Ng TK (2006) Rain removal in video by combining temporal and chromatic properties. In: ICME, pp 461–464, https://doi.org/10.1109/ICME.2006.262572
35. Zhang Y, Wang L, Qi J, Wang D, Feng M, Lu H (2018) Structured siamese network for real-time visual tracking. In: ECCV, pp 355–370, https://doi.org/10.1007/978-3-030-01240-3_22