

Image Deraining for UAV Using Split Attention Based Recursive Network

FENG Yidan, DENG Sen, WEI Mingqiang*

College of Computer Science and Technology/College of Artificial Intelligence, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, P.R. China

(Received 15 June 2020; revised 10 July 2020; accepted 25 July 2020)

Abstract: Images captured in rainy days suffer from noticeable degradation of scene visibility. Unmanned aerial vehicles (UAVs), as important outdoor image acquisition systems, demand a proper rain removal algorithm to improve visual perception quality of captured images as well as the performance of many subsequent computer vision applications. To deal with rain streaks of different sizes and directions, this paper proposes to employ convolutional kernels of different sizes in a multi-path structure. Split attention is leveraged to enable communication across multiscale paths at feature level, which allows adaptive receptive field to tackle complex situations. We incorporate the multi-path convolution and the split attention operation into the basic residual block without increasing the channels of feature maps. Moreover, every block in our network is unfolded four times to compress the network volume without sacrificing the deraining performance. The performance on various benchmark datasets demonstrates that our method outperforms state-of-the-art deraining algorithms in both numerical and qualitative comparisons.

Key words: unmanned aerial vehicle (UAV); deep neural network; image deraining; recursive computation; split attention

CLC number: TN925

Document code: A

Article ID: 1005-1120(2020)04-0539-11

0 Introduction

Unmanned aerial vehicles (UAVs) have a broad range of applications, such as disaster relief, aerial surveillance, film making, cargo transport and military reconnaissance. As an important outdoor imaging system, images captured by UAVs inevitably suffer from degradations caused by bad weather conditions. Rain, as the commonest bad weather, enormously decreases the visibility of the background image (Fig.1(a)), which has negative impacts on not only human perception but also subsequent computer vision applications.

Considering the forms of source inputs, rain removal algorithms can be divided into single image-based and video-based methods. Rain removal from videos, which are composed of sequential image frames, takes advantage of both the similarity in ad-

acent frames and the dynamics of rain to recover the background image. In contrast, single-image based deraining is more challenging due to the lack of temporal information. Early work solves this ill-posed problem mainly based on an optimization framework constrained by various image priors, such as Gaussian mixture model^[1] and low-rank representations^[2-4]. Limited by the manually designed models, they are only capable to remove certain types of rain streaks.



(a) Rainy image taken by UAV (b) UAV in rainy day

Fig.1 Image captured by UAV in rainy day and UAV affected by rain streaks

*Corresponding author, E-mail address: mqwei@nuaa.edu.cn.

How to cite this article: FENG Yidan, DENG Sen, WEI Mingqiang. Image deraining for UAV using split attention based recursive network[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2020, 37(4):539-549.

<http://dx.doi.org/10.16356/j.1005-1120.2020.04.005>

Recently, the performance of single image deraining is boosted by deep neural networks learning from massive data simulated using synthetic rain streaks of different directions and densities. For these methods, they usually train an end-to-end deep network model to learn a mapping between rainy images and corresponding ground truths that are regarded as no-rain images. In Ref.[5], a novel large-scale dataset containing 14 000 rainy/clean image pairs with 14 orientations and scales is synthesized, and then the ResNet model is conducted as the parameter layers to learn a negative mapping between clean and rainy images. In Ref.[6], a light-weight neuron attention (NA) architectural mechanism is presented to improve the channel attention mechanism, which adaptively recalibrates neuron-wise feature responses by modelling interdependencies and mutual influence between neurons. Stage-level information is concatenated and fused dynamically by NA module, which benefits rain removal of different types. The efforts of adopting and ameliorating advancing network designs for image deraining improve the learning ability and adaptability of deraining algorithms.

In this paper, we focus on the rain removal of images captured by UAV devices. Photography of UAVs is of high dynamicity and broad view. UAVs always feedback in the form of videos. High dynamicity renders the background changing more intensively in the same time interval, which greatly increases the difficulty in utilizing the temporal information. This difficulty can be alleviated by adopting the single-image-based rain removal algorithm frame-by-frame, but requires an efficient and light-weight solution to process rainy frames across different devices. Besides, UAVs take a bird's-eye view of the landscape, thus has a broad view on a variety of objects that are of abundant textures and scales. This requires a scale-aware deraining scheme to accurately separate rain streaks from the background image.

To this end, we propose a recursive multiscale split attention deraining network (RMSA-derainNet) for UAVs, where multiscale features are extracted in parallel branches and fused adaptively by

the split attention (SA) module in each basic block. Different from NASNet^[6], which uses channel-wise attention only to fuse stage-level features, we model feature-level interdependencies and mutual influences using SA. This separate-and-fuse process repeats several times (depending on the block number) in the deraining process. With only a few parameters increasing compared to a basic residual block, our RMSA-derainNet provides a light-weight but effective solution to tackle various types of rain streaks without spoiling important image details. Moreover, since modifications are made on block level, it is able to apply recursive computation among blocks, which leads to a much smaller model size. Extensive experiments show that our RMSA-derainNet is computationally efficient and addresses the deraining problem effectively for UAVs.

In summary, our contributions are as follows:

(1) We study the difficulty of rain removal on UAV images and propose a novel RMSA-derainNet for rain removal of UAV images. Grouping strategies are adopted on different levels of the network, which allows a light-weight (grouping of blocks) but effective (grouping of feature maps and their attentions) model for extracting fine-grained features and recovering detailed background information.

(2) SA and channel grouping strategy^[7-8] have been proved to be powerful tools for high level vision tasks such as recognition and classification. However, we for the first time demonstrate that these strategies combined with multi-scale convolutions can also boost the performance of the low level task on rainy image recovery, thanks to the adaptive receptive field and efficient information propagation.

(3) Extensive experiments demonstrate that our method can remove different types of rain streaks better and is able to recover fine background details, which not only benefits visual perception but also subsequent applications.

1 Related Work

In real applications, removing rain streaks

from a single image is an extremely challenging task for its ill-posed nature and lack of temporal information. To solve this issue, many deep-learning-based models have been proposed recently, which aim to extract the suitable low-level feature representation regarded as rain streaks. These methods generally contain a main structure which plays an important role in their architectures and preserve precise low-level feature representation during the training process.

Fu et al.^[5] first proposed to adopt a three-layer CNN to learn negative rain streak residuals from the high frequency components. In NASNet^[6], a lightweight NA architectural mechanism is presented to adaptively recalibrate neuron-wise feature responses by modelling interdependencies and mutual influence between neurons. Stage-level information is concatenated and fused dynamically by NA module, which benefits rain removal of different types. Zhang et al.^[9] labeled rain images with density levels to deal with rain streaks of different densities. Yang et al.^[10] learned the binary map of rain streak position to focus on rain affected areas. Wang et al.^[11] regularized the distance between the deraining result and the ground-truth image in the gradient domain. Wang et al.^[12] first learned the motion blur kernel of rain streak interference to guided the deraining network. Besides, Wang et al.^[13], Pan et al.^[14] and Yang et al.^[15] applied residual blocks to deepen the network. Specifically, Pan et al.^[14] proposed to recover image structures and details separately using two parallel network branches. Some deraining methods^[16] removed rain streaks progressively in multiple cascaded stages. In RESCAN^[16], different stages are connected using recurrent neural networks (RNN), which makes the information from previous stages available for the next stage. Li et al.^[17] employed dense blocks to fully connect different layers. Generative adversarial network (GAN)^[18] is exploited to refine the deraining results for more visual appealing effects. Li et al.^[19] put forth a two-stage network to first remove dense rain accumulation and then correct artefacts introduced by the first stage using a depth-guided GAN. Zhang et al.^[20] proposed a conditional GAN-based framework that consists of a

densely-connected generator and a multi-scale discriminator. There are some researches exploiting multi-branch or multi-stage network architecture for image rain removal. Parallel dual-branch structure is also proposed to learn different image components and ease the task in the single network branch. Deng et al.^[21] took advantage of another network branch to retrieve lost details, which validated the effectiveness of separate learning for the deraining problem.

In addition, there are some researches dedicated to improve the time and memory efficiency for increasing the practical value of rain removal algorithms. Fan et al.^[22] built a separable network supervised on block level, where a small number of blocks can be applied in resource-limited scenarios. Fu et al.^[23] simplified the learning process using shallow networks on the Laplacian image pyramid. Moreover, they proposed to utilize recursive computation among blocks to further reduce the number of parameters with negligible performance loss. The recursive strategy is also leveraged by PreNet^[24], which shares parameters not only among blocks but also among stages.

In general, existing researches either incorporate advancing network designs into the rain removal problem or combine problem-related information to guide the feature learning process. Popular network modules include residual block^[5,24], dense block^[15-17] or novel convolutional operations such as the dilated convolution^[9,19] and the squeeze-and-excitation^[15,19]. These modules have been proved to be not only effective for the high-level tasks, such as image recognition^[25] and object detection^[26], but also effective for the low-level tasks, e.g., single image deraining, image inpainting^[27], image colorization^[28] and image super resolution^[29]. In this paper, we present a new multi-scale attentive residual block as the basic modules of our network.

2 RMSA Derain Net

We propose an end-to-end neural network for single image rain removal, called RMSA-derain-Net. Multiscale paths are adopted to allow different levels of receptive field, which benefits the extrac-

tion of multiscale image semantics and the recovery of complex background. The overall framework is illustrated in Fig.2, our proposed network consists of novel multiscale attentive residual blocks (MARBs), which first combines multiscale paths with SA in a block structure, as shown in Fig.3. Particularly, in order to reduce the computational

complexity and memory cost of the network, we also implement the basic network function by recursively unfolding a group of MARBs several times, which can significantly reduce the number of parameters with negligible performance sacrifice. In the following sections, we will introduce MARB and the recursive network architecture in detail.

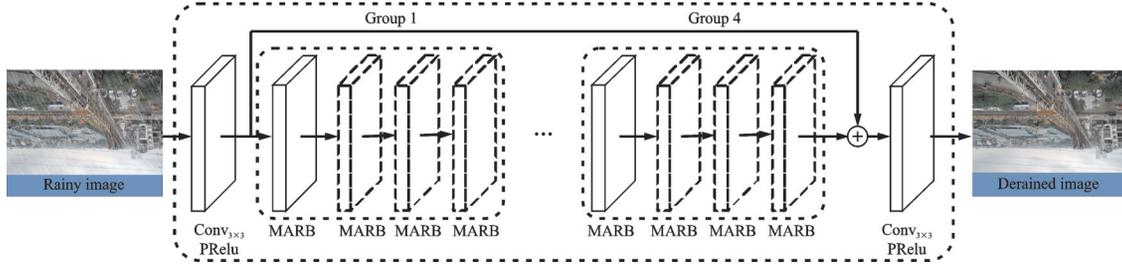


Fig.2 Overall framework of our proposed network

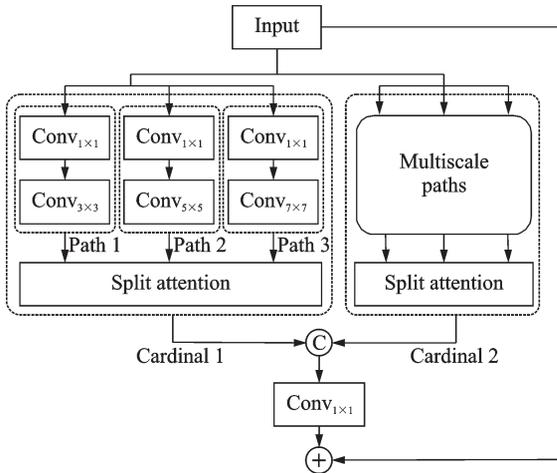


Fig.3 Multiscale attentive residual block

2.1 Multiscale attentive residual block

Due to their simple and modular structure, ResNet and its variants are widely used as the backbone network in image-to-image translation or other pixel-wise prediction tasks. In spite of the promising performances in different tasks, variants of ResNet suffer from massive parameters and require more computing resource. Inspired by the work in Ref.[7], we propose a MARB to construct our network for single image deraining, which only requires few more parameters than the original residual block but still achieves state-of-the-art performance.

As shown in Fig.3, the MARB used in our net-

work consists of two cardinal groups, and each cardinal group also contains three split paths that constructed by convolutional kernels in different scales. Multiscale kernels have different receptive fields and suit the need of filtering semantical objects in different scales. Each path consists of a 1×1 and a $k \times k$ convolutional layers. Denote the input and output feature of MARB as X and Y . Mathematically, the output Y can be described as

$$Y = \text{Add}(X, \text{Conv}_{1 \times 1}(\text{Cat}(\varphi_1(X), \varphi_2(X)))) \quad (1)$$

where $\varphi(\cdot)$ denotes the function of the cardinal group, which can be formulated as

$$\varphi(x) = \text{SA}(h_1(X), h_2(X), h_3(X)) \quad (2)$$

where $h(\cdot)$ indicates the function of the split path and $\text{SA}(\cdot)$ represents the function of the SA module (SAM). Specifically, $h(\cdot)$ in our network can be expressed as

$$h(x) = \text{Conv}_{1 \times 1}(\text{Conv}_{k \times k}) \quad (3)$$

where k in three split paths are set as 3, 5 and 7, respectively.

The detailed structure of SAM is illustrated in Fig.4. It is proposed in Ref.[30] as an enhancement of the original squeeze-and-excitation module. In SAM, features from different paths are added together and go through a normal “squeeze” process, which contains a global pooling layer and a fully connected (FC) layer that followed by the bath normal-

ization(BN) and the rectified linear unit(ReLU) activation function. Then, the features are split again to be fed into multiple FC paths and activated by r-Softmax function, thus “excite” each individual path using shared features. We argue that this particular structure of SAM can well model the interdependency and mutual influence among multiple paths. The reason of adopting SAM in the proposed multi-path structure is twofold: (1) SAM learns the weights of each path adaptively at training phase, therefore allows a flexible receptive field according to the source input, to adapt to the high dynamicity and broad view of UAV captured images with a better background objects recover in varying sizes; (2) SAM can be adopted at feature level rather than stage or branch level, which leads to more fine-grained feature learning. Feature-level SA can be incorporated into the block structure, where the original feature channels can be split to form multiple paths and thus make full use of information in limited channels.

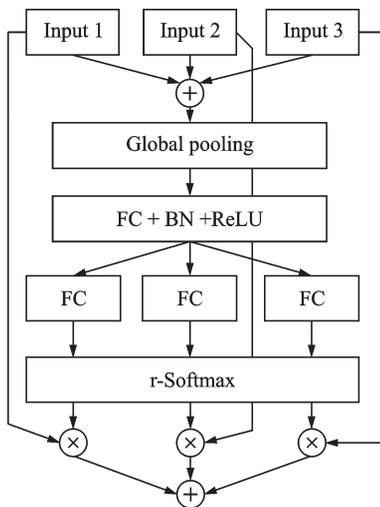


Fig.4 Split attention module^[7]

2.2 Recursive network architecture

The overall recursive network architectures are shown in Fig.2. The number of filters is set to be 64, and all the filters in our network are of size 3×3 and padding 1×1 . Our proposed network includes two convolution layers and 16 MARBs. The first layer can be interpreted as an encoder, which is

used to transform the rainy image into feature maps, and the last layer is used to recover the RGB channels from feature maps. Considering that the long skip connection compensates the long-range information and avoids gradient vanish, the feature maps of the first convolution layer are added to the last feature maps of MARB.

Moreover, motivated by Refs.[31-32], we divide 16 MARBs into four groups, each group consists of four MARBs. Then, we recursively unfold one MARB three times in each group so that every MARB in the same group shares the same parameters, as shown in Fig.5. Since network parameters mainly come from MARB, this intra-stage recursive design leads to a much smaller model size.

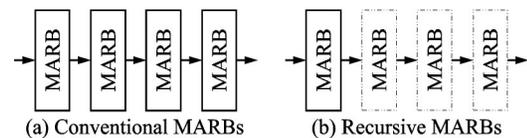


Fig.5 MARBs in one group

2.3 Comprehensive multi-scale loss function

Recently, hybrid loss functions, e.g., adversarial loss^[20], mean-square error (MSE) + structural similarity (SSIM)^[33], have been widely adopted for low level image tasks. Inspired by Ref.[34], we utilize a comprehensive multi-scale loss function L_c to train our network, which can be formulated as

$$L_c = \lambda_1 L_1 + \lambda_2 L_{MS-SSIM} \quad (4)$$

where λ_1 and λ_2 are the weighting parameters, which in our experiments are all fixed to be 1, and L_1 indicates the loss function, which can be formulated as

$$L_1 = \|f(I) - I'\| \quad (5)$$

where $f(\cdot)$ is the function that we try to learn, I the rainy image and I' the ground-truth rain-free image. Correspondingly, $L_{MS-SSIM}$ denotes the multiscale SSIM loss^[34], which calculates the SSIM value in different scales of image. Mathematically, it can be expressed as

$$L_{MS-SSIM} = 1 - MS_SSIM(I, I') \quad (6)$$

where $MS_SSIM(\cdot)$ denotes the operation of calcu-

lating different scales of SSIM. The combination of L_1 loss and multi-scale SSIM loss benefits the preservation of fine details in complex background, since L_1 loss tends to remain sharp edges and multi-scale SSIM loss constrains the structure similarity of the whole image at different scales.

3 Experiments

Our proposed method has been implemented on the Pytorch framework. For optimizing our network, the Adam is adopted with a min-batch size of 1 to train the network, where the size of image patch is set to $256 \text{ pixel} \times 256 \text{ pixel}$. We initialize the learning rate as 0.01, which is reduce by 50% every 30 epochs. For fully training our proposed network, we set the total number of training epochs to be 240, which takes around 1.5 d to train a model using the training set of Rain200H datasets. All the experiments are performed using a Nvidia 2080 GPU.

To validate the effectiveness of deraining models on UAVs captured images, we collect 500 UAVs captured no-rain outdoor images and 100 UAVs captured rainy images from Google and Youtube, and thus build a dataset containing both synthetic and real UAV rainy images. For synthetic images, we add different types of rain streaks on no-rain images using PS. This dataset is used to test and compare the performance of our method and existing deraining models.

3.1 Comparison with state-of-the-art methods

In this section, we compare our method with several state-of-the-art deraining methods on serval datasets, including two traditional deraining methods, i. e., Gaussian mixed model (GMM)^[1], discriminative sparse coding (DSC)^[4], and three learning-based deraining methods, i. e., deep detail network (DDN)^[5], spatial attentive network (SPA-Net)^[13], and dual CNN (DualCNN)^[14]. The codes of the compared methods are provided by their authors, which can be download from the Github. We

use the default configuration provided by the authors and only change the training datasets to train DDN, DualCNN and SPA-Net.

For the synthetic datasets, we evaluate the performance of our proposed method on three datasets, including Rain200H, Rain200L and Rain800. The Rain200H and Rain200L are provided by Yang et al.^[10], which contains 1 800 images for training and 200 images for testing. Compared with Rain200L, the Rain200H is more challenging, which is synthesized with five streak directions and the rain is more heavy. Zhang et al.^[20] also collected and synthesized the Rain800 dataset, which contains 700 training images and 100 testing images. Here we use peak signal to noise ratio (PSNR) and SSIM as evaluation metrics for the quantitative comparison. The quantitative evaluation results of PSNR and SSIM are shown in Table 1. As can be observed in Table 1, our proposed method obtains the highest values of PSNR and SSIM than other methods on three synthetic datasets.

Table 1 Quantitative experiments evaluated on three recognized synthetic datasets

Dataset	Rain200L		Rain200H		Rain800	
	PSNR	SSIM	PSNR	SSIM	PNSR	SSIM
GMM ^[1]	27.16	0.898 2	13.04	0.467 3	24.04	0.867 5
DSC ^[4]	25.68	0.875 1	13.17	0.427 2	20.95	0.753 0
DDN ^[5]	33.01	0.969 2	24.64	0.848 9	24.04	0.862 6
SPA-Net ^[13]	31.59	0.965 2	23.04	0.852 2	22.41	0.838 2
DualCNN ^[14]	32.93	0.957 5	24.09	0.763 2	23.83	0.839 5
Ours	35.27	0.972 9	26.95	0.889 2	25.88	0.889 5

The more visual comparisons are shown in Figs. 6—7, from which one can observe that our method better remains the structure and preserves the details of the images. For example, GMM and DSC cannot remove the rain streaks completely. Other deraining methods, i. e., DDN, DualCNN and SPA-Net remove the rain streaks but produce stripes. In the contrary, our method can remove the rain streak well.

For the real-world rainy datasets, since no rain-free ground truths for real-world images are provided, we perform a user study on several real-world

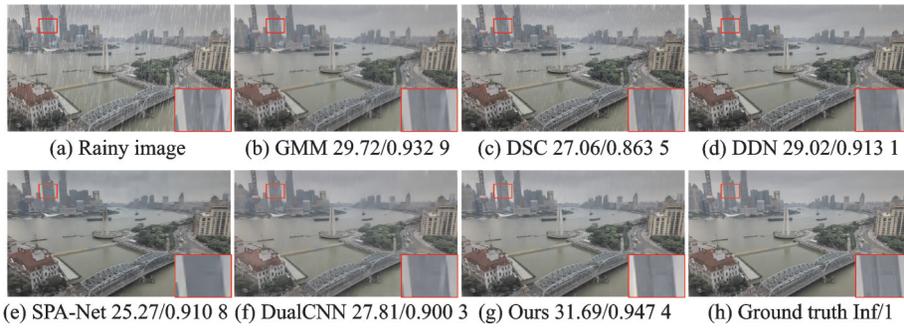


Fig.6 The first group of image deraining results tested in synthetic datasets using GMM, DSC, DDN, SPA-Net, DualCNN, and Ours, respectively

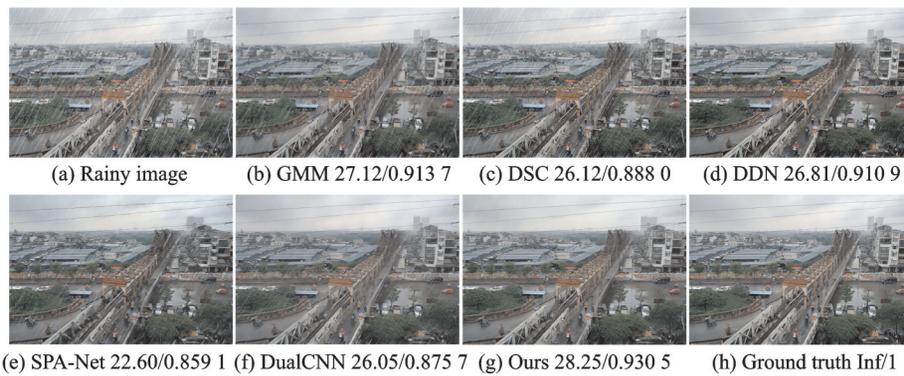


Fig.7 The second group of image deraining results tested in synthetic datasets using GMM, DSC, DDN, SPA-Net, DualCNN, and Ours, respectively

datasets collected by Yang et al.^[10] and Li et al.^[19].

Furthermore, to validate the practicability of our method, we also visually show the deraining re-

sults on some real-world rainy images in Fig.8. One can observe that our deraining result maintains more image details than others.

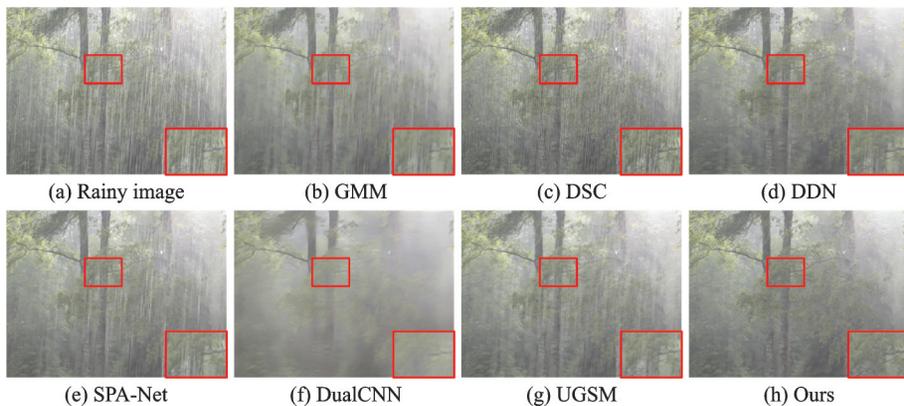


Fig.8 Image deraining results tested in real-world datasets using GMM, DSC, DDN, SPA-Net, DualCNN, UGSM^[35] and Ours, respectively

Inspired by Fu et al.^[36], we randomly select 30 rainy images from the real-world rainy datasets. Secondly, we utilize the state-of-the-art methods to remove the rain streaks from the above images. Then, 10 people are asked to score the derained results along with its original rainy image from 1 to 5

(1 represents the worst quality and 5 denotes the best quality), which are randomly presented without knowing the corresponding methods.

The average scores of different deraining methods are shown in Fig.9, where each point in the scatter plot indicates an image. The horizontal axis

denotes the average score of rainy images, and the vertical represents the average score of derained im-

ages. From Fig.9 one can see that, our method also achieves a better result.

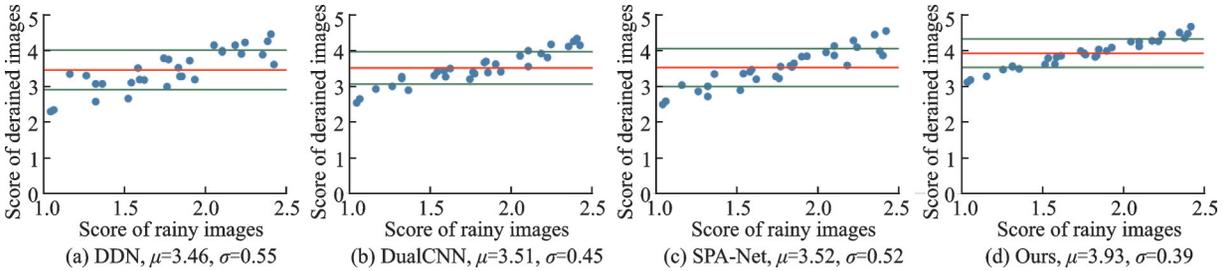


Fig.9 Average scores of DDN, DualCNN, SPA-Net, and Ours

3.2 Ablations and discussion

In this subsection, we assess the effect of several key modules of our network, including MARB and the recursive computation. All the ablation studies are conducted on Rain200H, which contains 1 800 rainy images for training and 200 rainy images for testing.

Here we replace MARB with residual block (RB) in our network and show the result in Table 2. One can see that the network with MARB leads to higher average PSNR and SSIM values, which indicates the effectiveness of the proposed MARB.

Table 2 Results of our network by replacing MARB with RB

Dataset	Metric	RB	MARB
Rain200H	PSNR	26.20	26.95
	SSIM	0.875 6	0.889 2

Our proposed network adopts recursive computation, which can reduce the parameters of the network. In Table 3, we evaluate the deraining result with and without recursive computation in our network, and find negligible loss in the deraining performance.

Table 3 Performance and the number of parameters with and without recursive computation in our network

Metric	Recursive	Without recursive
PSNR	26.95	27.86
SSIM	0.889 2	0.915 2
Parameter number	78 153	312 227

4 Conclusions

This paper proposes a deraining method for the recovery of UAV captured images using SA based recursive network. A novel MARB is constructed by enabling attention on different groups of feature maps, which are obtained after conducting multi-scale convolution operations. The learned feature representations are universally improved to a boost the performance of single image deraining, which is validated by our ablation study in Table 2. With this improvement, image structures and semantically important details are well preserved after removing the rain streaks, leading to a better value in use.

Furthermore, the recursive computation is adopted at module level, which allows several blocks collaborate to remove rain streaks progressively using the same parameters. In Table 3, we demonstrate that the model size is greatly compacted with only negligible sacrifice in deraining performance, which increases the practical value for rainy image recovery in resource-limited platforms.

In the future work, since rainy haze caused by rain drop accumulation can be easily observed under the UAV's view, efforts can be made to remove not only rain streaks but also rainy haze while keeping the deraining model compact and efficient for real-time and cross-platform use.

References

- [1] LI Y, TAN R T, GUO X, et al. Rain streak removal using layer priors [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

- Las Vegas, NV, USA: IEEE, 2016: 2736-2744.
- [2] ZHANG H, PATEL V M. Convolutional sparse and low-rank coding-based rain streak removal[C]//2017 IEEE Winter Conference on Applications of Computer Vision (WACV). [S.l.]: IEEE, 2017: 1259-1267.
- [3] CHEN Y L, HSU C T. A generalized low-rank appearance model for spatio-temporally correlated rain streaks [C]//Proceedings of the IEEE International Conference on Computer Vision. Sydney, Australia: IEEE, 2013: 1968-1975.
- [4] LUO Y, XU Y, JI H. Removing rain from a single image via discriminative sparse coding[C]//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 3397-3405.
- [5] FU X, HUANG J, ZENG D, et al. Removing rain from single images via a deep detail network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE, 2017: 3855-3863.
- [6] QIN X, WANG Z. NASNet: A neuron attention stage-by-stage net for single image deraining[EB/OL]. (2019-12-06) [2020-05-30]. <http://arXiv preprint arXiv:1912.03151>.
- [7] ZHANG H, WU C, ZHANG Z, et al. ResNeSt: Split-attention networks[EB/OL]. (2020-04-19) [2020-05-30]. <http://arXiv preprint arXiv:2004.08955>.
- [8] LI X, WANG W, HU X, et al. Selective kernel networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019: 510-519.
- [9] ZHANG H, PATEL V M. Density-aware single image de-raining using a multi-stream dense network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 695-704.
- [10] YANG W, TAN R T, FENG J, et al. Deep joint rain detection and removal from a single image[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE, 2017: 1357-1366.
- [11] WANG Y, ZHANG H, LIU Y, et al. Gradient information guided deraining with a novel network and adversarial training[EB/OL]. (2019-10-09) [2020-05-30]. <http://arXiv preprint arXiv:1910.03839>.
- [12] WANG Y T, ZHAO X L, JIANG T X, et al. Rain streak removal for single image via kernel guided CNN [EB/OL]. (2018-08-26) [2020-05-30]. <http://arXiv preprint arXiv:1808.08545>.
- [13] WANG T, YANG X, XU K, et al. Spatial attentive single-image deraining with a high quality real rain dataset[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019: 12270-12279.
- [14] PAN J, LIU S, SUN D, et al. Learning dual convolutional neural networks for low-level vision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 3070-3079.
- [15] YANG Y, LU H. Single image deraining using a recurrent multi-scale aggregation and enhancement network [C]//2019 IEEE International Conference on Multimedia and Expo (ICME). Shanghai, China: IEEE, 2019: 1378-1383.
- [16] LI X, WU J, LIN Z, et al. Recurrent squeeze-and-excitation context aggregation net for single image deraining[C]//Proceedings of the European Conference on Computer Vision (ECCV). [S.l.]: Springer, 2018: 254-269.
- [17] LI G, HE X, ZHANG W, et al. Non-locally enhanced encoder-decoder network for single image deraining [C]//Proceedings of the 26th ACM International Conference on Multimedia. Daejeon, Korea: ACM, 2018: 1056-1064.
- [18] JIN X, CHEN Z, LIN J, et al. AI-GAN: Signal de-interference via asynchronous interactive generative adversarial network[C]//Proceedings of 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). [S.l.]: IEEE, 2019: 228-233.
- [19] LI R, CHEONG L F, TAN R T. Heavy rain image restoration: Integrating physics model and conditional adversarial learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019: 1633-1642.
- [20] ZHANG H, SINDAGI V, PATEL V M. Image deraining using a conditional generative adversarial network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2019. DOI: <http://10.1109/TCSVT.2019.2920407>.
- [21] DENG S, WEI M, WANG J, et al. DRD-Net: Detail-recovery image deraining via context aggregation net-

- works[EB/OL]. (2019-08-28) [2020-05-30]. <http://arXiv preprint arXiv:1908.10267>.
- [22] FAN Z, WU H, FU X, et al. Residual-guide network for single image deraining [C]//Proceedings of the 26th ACM International Conference on Multimedia. Daejeon, Korea: ACM, 2018: 1751-1759.
- [23] FU X, LIANG B, HUANG Y, et al. Lightweight pyramid networks for image deraining [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2019, 31(6): 1794-1807.
- [24] REN D, ZUO W, HU Q, et al. Progressive image deraining networks: A better and simpler baseline [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA: IEEE, 2019: 3937-3946.
- [25] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04) [2020-05-30]. <http://arXiv preprint arXiv:1409.1556>.
- [26] REN S, HE K, GIRSHICK R, et al. Faster RCNN: Towards real-time object detection with region proposal networks [J]. *IEEE Trans Pattern Anal Mach Intell*, 2017, 39(6): 1137-1149.
- [27] LI H, LI G, LIN L, et al. Context-aware semantic inpainting [J]. *IEEE Transactions on Cybernetics*, 2018, 49(12): 4398-4411.
- [28] ZHANG W, FANG C W, LI G B. Automatic colorization with improved spatial coherence and boundary localization [J]. *Journal of Computer Science and Technology*, 2017, 32(3): 494-506.
- [29] LAI W S, HUANG J B, AHUJA N, et al. Deep laplacian pyramid networks for fast and accurate super-resolution [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE, 2017: 624-632.
- [30] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [31] KIM J, KWON LEE J, MU LEE K. Deeply-recursive convolutional network for image super-resolution [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA: IEEE, 2016: 1637-1645.
- [32] TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE, 2017: 3147-3155.
- [33] FAN Z, WU H, FU X, et al. Residual-guide feature fusion network for single image deraining [EB/OL]. (2018-04-20) [2020-05-30]. <http://arXiv preprint arXiv:1804.07493>.
- [34] ZHAO H, GALLO O, FROSIO I, et al. Loss functions for neural networks for image processing [EB/OL]. (2015-11-28) [2020-05-30]. <http://arXiv preprint arXiv:1511.08861>.
- [35] DENG L J, HUANG T Z, ZHAO X L, et al. A directional global sparse model for single image rain removal [J]. *Applied Mathematical Modelling*, 2018, 59: 662-679.
- [36] FU X, QI Q, HUANG Y, et al. A deep tree-structured fusion model for single image deraining [EB/OL]. (2018-11-21) [2020-05-30]. <http://arXiv preprint arXiv:1811.08632>.
- Acknowledgement** The work was supported by the Foundation of Graduate Innovation Center in Nanjing University of Aeronautics and Astronautics (No. kfjj20191601).
- Authors** Ms. FENG Yidan received the B.S. degree from the Department of Measurement and Control Technology, Hefei University of Technology, Hefei, China, in 2017. She is currently pursuing the M.S. degree with Nanjing University of Aeronautics and Astronautics, Nanjing, China. Her research interests focus on computer vision and machine learning.
- Dr. WEI Mingqiang received his Ph.D. degree (2014) in Computer Science and Engineering from the Chinese University of Hong Kong (CUHK). He is an associate professor at the School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics (NUAA). Before joining NUAA, he served as an assistant professor at Hefei University of Technology, and a postdoctoral fellow at CUHK. His research interest is computer graphics with an emphasis on geometry processing and image analysis.
- Author contributions** Ms. FENG Yidan conducted the analysis and wrote the manuscript. Mr. DENG Sen contributed to data and the code of the model. Dr. WEI Mingqiang designed the study, revised and modified the

manuscript. All authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing interests.

(Production Editor: ZHANG Huangqun)

基于分散注意力递归神经网络的无人机去雨算法

冯一箖, 邓 森, 魏明强

(南京航空航天大学计算机科学与技术学院/人工智能学院, 南京 211106, 中国)

摘要: 户外拍摄的图像常受降雨天气影响, 造成图像可见度低等问题。作为一种重要的户外图像采集系统, 无人机需要合适的去雨算法来提高雨天所摄图像的视觉效果, 并给后续计算机视觉任务提供高质量的初始图像数据。为处理不同尺度和方向的雨纹, 本文提出一种利用和融合多尺度卷积的多分支网络结构, 并与分散注意力机制结合, 实现特征级别的多通道交互, 从而利用自适应的感受野以处理复杂背景的恢复问题。本文算法将多尺度分支、分散注意力机制和通道分组机制相结合, 在实现多尺度处理的基础上与基本残差块的特征通道数保持一致。另外, 为进一步减少网络模型的参数量, 本文算法将每个基本模块递归计算4次, 从而在保持去雨效果的同时使模型轻量化。大量实验证明, 本文算法在多个数据集上均优于目前主流的传统或基于学习的图像去雨方法。

关键词: 无人机; 深度神经网络; 单幅图像去雨; 递归计算; 分散注意力