IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 7, 2021, accepted April 2, 2021, date of publication April 14, 2021, date of current version April 22, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3073127

Enhanced Attentive Generative Adversarial Network for Single-Image Deraining

GUOQIANG CHAI[1,2,3], ZHAOBA WANG[1,3], GUODONG GUO[1], (Senior Member, IEEE), YOUXING CHEN[1], YONG JIN[1], DAWEI WANG[2], BIN LU[2], AND SHILEI REN[1]

[1]School of Information and Communication Engineering, North University of China, Taiyuan 030051, China
[2]School of Physics and Information Engineering, Shanxi Normal University, Linfen 041000, China
[3]National Key Laboratory for Electronic Measurement Technology, North University of China, Taiyuan 030051, China

Corresponding author: Zhaoba Wang (wangzb@nuc.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62004119, in part by the Natural Science Foundation of Shanxi Province under Grant 201801D121150 and Grant 201901D111155, and in part by the Applied Basic Research Project of Shanxi Province under Grant 201801D121160.

ABSTRACT The problem of image rain removal has drawn widespread attention as the blurry images caused by rain streaks can degrade the performance of many computer vision tasks. Although exiting deep learning-based methods outperform most traditional methods, there are still unresolved issues in terms of performance. In this paper, we propose a novel enhanced attentive generative adversarial network named EAGAN to effectively remove the rain streaks and restore the image structural details at the same time. As rain streaks have different sizes and shapes, EAGAN utilizes a multiscale aggregation attention module (MAAM) to produce an attention map to guide the subsequent network to put conscious attention to rain regions. A symmetrical autoencoder with long-range skip-connections, squeeze-and-excitation (SE) modules, and non-local operation is further utilized to enhance the representation of the network. Finally, spectral normalization and a relativistic generative adversarial network (GAN) are further applied to improve the training stability and deraining performance. Both qualitative and quantitative validations on synthetic and real-world datasets demonstrate that the proposed approach can achieve a competitive performance in comparison with the state-of-the-art methods.

INDEX TERMS Image deraining, attentive module, relativistic GAN.

I. INTRODUCTION

As a common weather condition, rain streaks not only bring pool visual perception for multimedia applications but also significantly degrade the performance of many computer vision algorithms, which are performed on high quality and reliable images, such as object tracking [1], semantic segmentation [2], and outdoor surveillance [3]. Under rainy weather, rain streaks with various orientations and scales accumulate and present undesirable impairment on the qualities of captured images. Therefore, it is necessary to explore effective rain removal methods before executing regular algorithms to ensure a high performance in images.

In the past decade, image deraining has drawn widespread attention and several approaches have been proposed to address the problem. Existing techniques can be roughly divided into two categories according to their nature, i.e., video-sequence based methods and single-image based methods. As the same background can be available in the successive frames, video sequence-based methods [4]–[9] attempt to handle the problem by exploiting the temporary correlation information and frequency properties of rain streaks. Due to the lack of temporal information, single-image based methods are more challenging than the video sequence-based. However, single-image based methods have attracted considerable research attention due to its greater practicality and time-efficiency.

For single image deraining, several traditional optimization-based algorithms, such as nonlocal means filter [10], sparse coding [11], low rank approximation [12], representation learning [13] and pre-trained Gaussian mixture models (GMM) [14] have been proposed to deal with the problem. These traditional methods attempt to explore certain prior information on texture characteristics to model the rain streaks and then separate them from the background [11]–[14], or directly restore the image with nonlocal

The associate editor coordinating the review of this manuscript and approving it for publication was Mohamad Forouzanfar.

mean smoothing [10]. However, these predefined models have a limited capability to represent the real rain streaks, resulting in that they are only effective for certain patter and fail to restore image details of the same nature as rain streaks. Furthermore, due to a large number of parameters that need to be optimized, these optimization-based algorithms are usually time-consuming.

Recently, deep learning-based methods have achieved a great success in many computer vision problems due to its powerful ability of high-level feature representation. There are also several deep learning-based methods [15]–[28] addressing the single-image rain removal problem by directly establishing an end-to-end projection between the rainy images and the clean ones.

Although these deep learning-based approaches have achieved better performance than traditional methods in terms of time-efficient and deraining quality, there are still two major challenges. One is that some methods [15]–[19] focus on the entrance of the networks and decompose the input image into low and high frequencies, which increases the complexity of the algorithm and may introduce an incorrect luminance information [17]. Another challenge is that the deraining performance needs to be improved. As rain streaks are embedded in the background and the scene covered by rain is seriously damaged, some methods tend to leave too many rain streaks in the background, while other methods tend to over-smooth the image details.

To avoid the risk of incorrect luminance information and improve the deraining performance, we propose an enhanced attentive generative adversarial network (EAGAN), which can remove rain streaks and restore the realistic scenes from the original image rather than the decomposed frequency. In EAGAN, a multiscale aggregation attention module (MAAM) is utilized to produce an attention map to guide the subsequent network to put conscious attention to rain regions, then the spatial contextual features extracted by non-local neural module (NLNM) along with the attention map is fed into a symmetrical autoencoder to produce the rain-free image. Finally, a relativistic discriminator is utilized to promote the generator to produce realistic scenes. In summary, this work makes the following contributions:

(1) A novel enhanced attentive generative adversarial network is proposed to deal with the single-image deraining problem, which directly establishes an end-to-end projection from the rainy images to the clean ones.

(2) To make the generator focus accurately on rain regions in the spatial dimension, we design a multiscale aggregation attention module with various convolution kernel sizes and self-adaptive residual manner to generate the rain attention map. A symmetrical autoencoder with long-range skip-connections, squeeze-and-excitation block, and region level non-local operation is applied for a better high-level feature representation.

(3) The attention map is used as additional supervised information for the discriminator to make it pay conscious attention to the rain streaks. Meanwhile, a relativistic GAN, rather than the conditional GAN, is applied to further improve the performance.

(4) To enhance the training stability and produce realistic images, spectral normalization is utilized both in the generator and discriminator. Experimental results on both synthetic and real-world datasets have demonstrated the effectiveness of the proposed EAGAN, which achieves superior performance compared to the state-of-the-art methods.

The remaining sections are organized as follows: In Section II, the related works on rain removal task are reviewed briefly. In Section III, the details of the proposed EAGAG, as well as the loss function are presented. Subsequently, experimental settings, qualitative and quantitative evaluations are presented in Section IV. Finally, the conclusion and discussion are given in Section V.

II. RELATED WORK

In this section, we present a brief review of rain removal methods in the past decade, including video sequence-based and single-image based ones.

A. VIDEO-SEQUENCE BASED METHODS

Since rich temporal information can be derived from successive image frames to provide additional information for rain streaks detection and removal, video sequence-based methods have been widely explored. Garg and Nayar [4], [5] were the pioneers to focus on rain removal issue, they assumed the background scene to be static, then detected rain streaks by comparing the intensity change with an empirical threshold, and finally removed them by averaging the intensity of rain pixels with the previous and subsequent information. Kim *et al.* [6] defined a rain map with the optical flow information in successive frames and then removed rain streaks by utilizing a low-rank matrix completion technique. Jiang *et al.* [7] combined total variation regularizer, tensor nuclear norm, and time directional difference operator to remove rain streaks directly without rain region detection. Li *et al.* [8] assumed the rain streaks with multiscale configurations sparsely scatted over different regions and removed them with multiple convolutional filters. Li *et al.* [9] proposed a rain/snow removal approach from surveillance videos, which fully considers dynamic statistics of both rain/snow and background scenes taken from a video sequence.

B. SINGLE-IMAGE BASED METHODS

Since the temporal information is unavailable, single-image deraining is more challenging compared to the video sequence-based. However, single-image deraining is more practical than video sequence-based one, as sometimes there are only images can be available in the daily life, rather than the video data. In addition, image processing is more time-efficiency than video processing.

Luo *et al.* [11] proposed a discriminative sparse coding approach based on a strong mutual exclusivity property to separate the rain layers from the background. Li *et al.* [14] formulated the rain removal as a layer decomposition problem

and proposed Gaussian mixture models (GMM) to recognize rain streaks with various orientations and scales. Chang *et al.* [20] incorporated a low-rank prior transformation to separate rain streaks from images based on the fact that the streaks have an extremely distinct low-rank structure. Zhu *et al.* [29] introduced three image priors that first identify rain regions by analyzing the local gradient statistics and then remove them by predefined priors on the background layer. Chen and Hsu [12] proposed a tensor structure model to detect the spatiotemporally correlated rain streaks and then remove them in a unified way.

In recent years, deep learning-based methods have been utilized in rain streaks removal due to its powerful high-level feature representation and superior ability to learn nonlinear functions. Yang and Lu [24] applied hierarchy enhancement units to fully extract local information and then utilized recurrent enhancement units to remove rain streaks stage by stage. Wang *et al.* [21] first provided a human supervision rain/clean image pairs for training and then proposed a residual network with a spatial attention mechanism to remove rain streaks in a local-to-global manner. Yang *et al.* [22], [41] proposed a multi-task architecture that first learns a binary rain streak map and then separates rain streaks with a contextualized dilated module, recently, they provided a comprehensive survey of deraining methods over the last decade [42]. Li *et al.* [23] adopted a modified dilated convolutional neural network to enlarge receptive field, and then remove rain streaks by incorporating squeeze-and-excitation (SE) [32] block and recurrent neural network. Zhang *et al.* [25] provided a widely used rain dataset and proposed a conditional generative adversarial network to deal with dearining issue. Qian *et al.* [26] synthesized a training dataset covered by raindrops and remove them from the background by an attention GAN with a multiscale objective function. Chai *et al.* [27], [46] proposed two recurrent networks to remove rain streaks stage by stage. Yasarla and Patel [28] adopted an uncertainty guided network to detect rain streaks and then remove them with cycle spinning enhancement. Zhang and Patel [38] proposed a multiscale dense network to estimate rain density and then remove them with another multiscale dense network. Du *et al.* [43] proposed a variational image deraining (VID) method by formulating image deraining in a conditional variational auto-encoder framework. Jiang *et al.* [44] proposed to decompose rain streaks into multiple rain layers and individually estimate each of them along the network stages to cope with the deraining problem. Jiang *et al.* [45] applied the multi-scale collaborative representation for rain streaks from the perspective of input image scales and hierarchical deep features in a unified framework to remove rain streaks.

All these deraining methods have achieved significant progresses in rain streaks removal, however, they tend to remain too many rain streaks or miss image details in the results, i.e., the deraining performance needs to be significantly improved.

III. PROPOSED METHOD

In this section, the attentive network architecture of EAGAN is presented, and then the spectral normalization relativistic GAN is demonstrated in detail. At last, the loss function for training EAGAN to restore realistic deraining scenes is presented.

A. NETWORK ARCHITECTURE

The overall architecture of the proposed EAGAN is illustrated in Fig. 1. It can be divided into two main parts, i.e., a generator and a discriminator.

The generator first applies a multiscale aggregation attention module (MAAM) to produce an attention map to guide the subsequent network to put conscious attention to rain regions. Compared with the multiscale aggregation module in [22] and [38] which is just used to extract features, our MAAM incorporates residual learning and integrates the relationships between channels into the feature map. The region level non-local neural module (NLNM) enhances the representation capability by exploiting the abundant structure cues and spatial self-similarity information. Then the feature maps generated by NLNM and the attention map are fused in a self-adaptive residual manner. Afterward, the deraining image is produced by a symmetrical squeeze-and-excitation autoencoder module. Furthermore, long-range skip-connections that bypass intermediate layers are applied to keep the delivery of image details from low-level to high-level.

The discriminator is utilized to differentiate the deraining results from the ground-truths. Instead of the standard GAN, we apply the relativistic GAN combined with an attention loss for a better deraining performance.

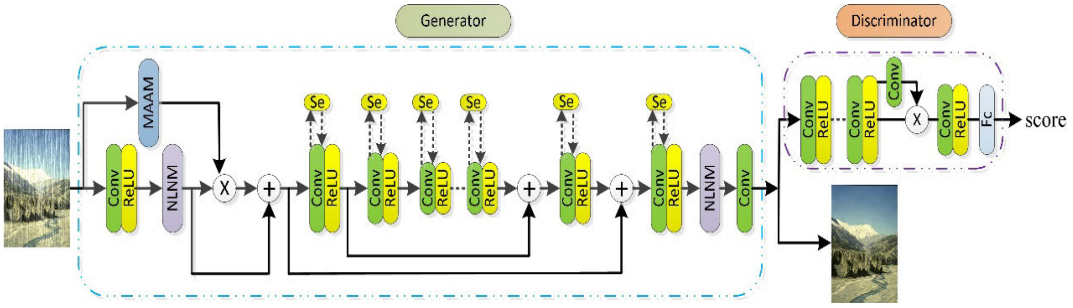


FIGURE 1. The architecture of proposed EAGAN.

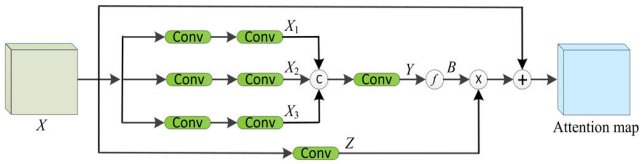


FIGURE 2. The schematic illustration of multiscale aggregation attention module.

B. GENERATIVE NETWORK

The generator first applies a multiscale aggregation attention module with various convolution kernel sizes and learnable residual to generate the rain attention map. A symmetrical autoencoder with long-range skip-connections, squeeze-and-excitation block, and region level non-local operation is applied for a better high-level feature representation.

1) MULTISCALE AGGREGATION ATTENTION MODULE

The visual attention mechanism has been widely applied in computer vision algorithms as its such excellent capability to locate targeted regions. Rainy images may contain different sizes and shapes of rain streaks, smaller rain streaks could be captured by smaller receptive fields while the larger ones could be captured by larger receptive fields. To this end, we design a multiscale aggregation attention module that combining features from different scales to locate the rain regions.

The schematic illustration of MAAM is presented in Fig. 2. The input feature map is first fed into parallel convolutional layers that consist of three channels with different filter sizes. Secondly, the parallel outputs with abundant contextual information are concatenated and fused further by a convolution with the kernel size of 1×1, the output of this convolution is a set of unnormalized attention weight $\{y_1, y_2, \ldots y_n\}$ and each weight represents a specific type of rain streaks, which can be formulated as

$$\begin{cases} X_i = F^{(k)}(X), \quad i \in \{1, 2, 3\},\ k \in \{1, 3, 5\} \\ Y = F^{(1)}([X_1, X_2, X_3]) \end{cases} \tag{1}$$

where X is the feature map generated from the preceding layers, X_i is the output of the i'th channel and $F^{(k)}(\cdot)$ represent the convolution with the kernel size of $k \times k$ followed by an activation function. Specifically, the kernel size of the three channels is 1, 3, and 5, respectively. $[X_1, X_2, X_3]$ represents the concatenation of feature maps generated by the multiscale convolution channels.

Furthermore, a softmax layer is applied to normalize the weights and the channel attention weights $\{b_1, b_2, \ldots b_n\}$ can be finally formulated as

$$b_{i,x} = \frac{\exp(y_{i,x})}{\sum_{c=1}^{n} \exp(y_{i,x})} \tag{2}$$

where $y_{i,x}$ is the output at pixel x of the i'th channel of the fused convolution layer and $b_{i,x}$ is the calculated attention weight.

Afterward, the attention weight is multiplied with a new feature map Z generated by another bypass in an elementwise manner. At last, the multiplied result is added to the input feature map X by a learnable manner, and the operation can be formulated as:

$$A = \lambda(B \cdot Z) + X \tag{3}$$

where λ is the learnable parameter, B is the fused feature maps after softmax, and A is the final attention map which is used for the subsequent network to put conscious attention to rain regions.

Fig. 3 shows a rainy image and its attention maps generated by different kernel sizes. The rain streaks are shorter in the upper image and longer in the lower image. As can be observed, the smaller kernel size can acquire the shorter rain streaks and other details but tend to blur the longer rain streaks. Conversely, the larger kernel size has a stronger ability to obtain the longer streaks. Therefore, the incorporation of different kernel size in MAAM can obtain both the small and large streaks as well as the details.

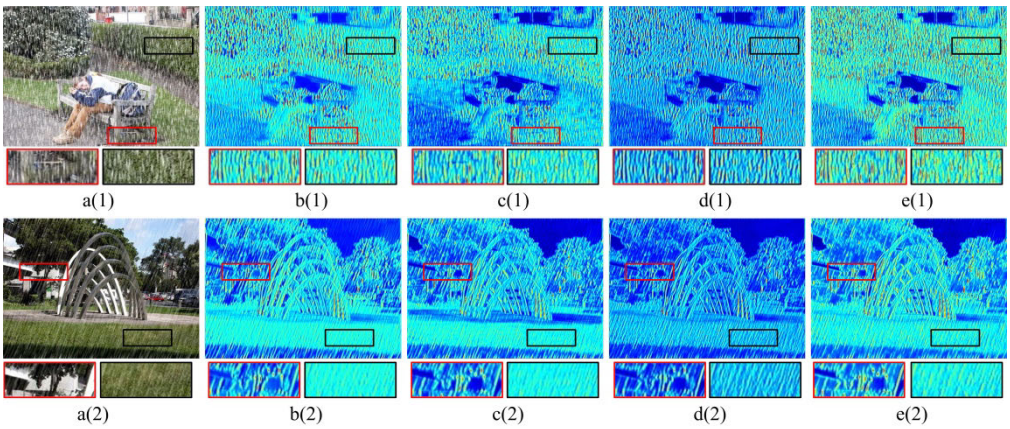


FIGURE 3. The attention maps generated by different channels in MAAM, the red areas indicate a high probability of rain. (a1-a2) Rainy image. (b1-b2) Results by ks = 1. (c1-c2) Results by ks = 3. (d1-d2) Results by ks = 5. (e1-e2) Results by MAAM.

2) NON-LOCAL NEURAL MODULE

Wang *et al.* [30] adopted convolutional neural networks (CNN) to realize non-local operation for the task of video classification. The non-local neural networks have an excellent capability to capture long-range spatial dependencies throughout the entire image, which is very useful for high-level computer vision tasks.

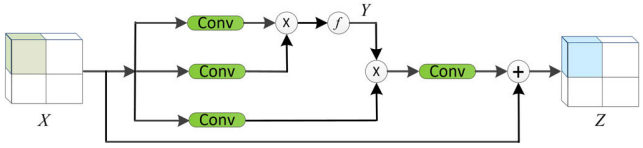


FIGURE 4. The schematic illustration of non-local neural module.

The schematic illustration of NLNM is presented in Fig. 4. We apply the extensional Gaussian function to compute the similarity in an embedded space [30], and the non-local operation can be formulated as

$$\begin{cases} Y = \text{softmax}[(F^{(1)}(X))^T \cdot F^{(1)}(X)] \\ Z = F^{(1)}[F^{(1)}(X) \cdot Y] + X \end{cases} \tag{4}$$

where X is the feature map generated from the preceding layers, $F^{(1)}(\cdot)$ represents the convolution with the kernel size of 1×1 followed by an activation function, and T represents the matrix transpose.

The traditional global-level non-local operations may incur an unacceptable computational burden, especially when the size of a feature map is large. Besides, Liu *et al.* [31] has demonstrated that non-local operations with an appropriate neighborhood size have more advantageous for low-level tasks, such as image deraining and super-resolution. Therefore, the non-local operations are executed at a region level in this work. To be specific, we first divide the input feature map into a grid of regions (As shown in Fig. 4, $k \times k$ indicates that the input feature map is divided into k^2 sub-regions with the same size), and then let the subsequent layers process the feature map inside each region.

As analyzed above, the capability of feature representation is enhanced by the region-level operation via exploiting the spatial self-similarities.

3) SYMMETRICAL AUTOENCODER MODULE

The purpose of the symmetrical autoencoder is to restore an image that is free from rain streaks and artifacts produced by the previous stage.

The proposed autoencoder contains 12 convolution-ReLU blocks and symmetrical skip-connections. The benefits of these skip-connections can be described in the following two aspects: first, they strengthen the feature propagation and make sure that the low-level information is still available in the very end of the whole architecture. Second, they alleviate the vanishing-gradient problem and strengthen the robustness of network.

To better model the interdependencies of the feature maps, we add a SE block to each convolution layer. Compared to the pure convolution operation which only considers the spatial information in each channel, SE adaptively recalibrates channel-wise feature responses by explicitly modeling interdependencies between channels [32]. Specifically, SE utilizes a self-gating learning mechanism based on the channel dependence to produce different channel weights which are further used to perform the recalibration of features.

Besides, a NLNM is applied after the autoencoder for a better representation and a convolution layer without any activation function is added at the end to generate the deraining image. As instance normalization (IN) and batch normalization (BN) may generate unpleasant artifacts [17], [27], we apply spectral normalization (SN) for all convolution layers to enhance the training stability while improving the deraining performance.

C. RELATIVISTIC DISCRIMINATOR NETWORK

Since it is usually the scene that covered by rain streaks suffers from information damage seriously, we want to make the discriminator pay more attention to these areas. To address this problem, we use the attention map produced by MAAM as the supervised information. Specifically, we extract the features from the interior layers and feed them to a mask convolution, then compute the pixel loss between the output and the supervised map. Furthermore, the output of mask convolution is utilized to recalibrate the features of the previous layer by multiplying them in an elementwise manner. Finally, a fully connected layer is used at the end layer to calculate the probability that an image is a real scene. Meanwhile, similar to the generator, we also apply SN for all the convolution layers in the discriminator to enhance the deraining performance.

In standard GAN, the discriminative network performs a binary classification to distinguish whether an input image is real or fake, and the generator is trained to generate a fake image that can convince the discriminator that it is real. While in relativistic GAN, the network is also trained to decrease the probability that the real image is real, and the discriminator tries to estimate the probability that a real sample is more realistic than the fake one.

We utilize the relativistic average discriminator (RaD) proposed in [33] to compare the realistic probability of the input image to the average realistic probability of the opposite image. The discriminator in a standard GAN can be described as $D(x) = \sigma(C(x))$, where x donates a real or fake image, $C(x)$ is the output of non-transformed layer and σ donates the sigmoid activation function. While in relativistic average discriminator, the output depends on both real and fake images, and it can be defined as $D_{Ra}(x_r, x_f) = \sigma(C(x_r) - E_{x_f}[C(x_f)])$, where $E_{x_f}[C(x_f)]$ donates the average of all the fake images in a mini-batch. The GAN loss for the discriminator can be expressed as:

$$\begin{aligned} L_{D\_Ra} = &-E_{x_r}[\log(D_{Ra}(x_r, x_f))] \\ &-E_{x_f}[\log(1 - D_{Ra}(x_f, x_r))] \end{aligned} \tag{5}$$

D. LOSS FUNCTION

To achieve a better deraining performance, we combine the content loss, negative structural similarity (SSIM) [34] loss, perceptual loss, attention loss, and GAN loss to train the proposed EAGAN. Specifically, the content loss, which can effectively evaluate the pixel-level differences between deraining result and its corresponding ground-truth with a Euclidean distance, can be expressed as:

$$L_c = E_x||G(x) - y||_2^2 \tag{6}$$

where x is the rainy image, G(x) is the deraining result generated by the proposed EAGAN, and y is the ground-truth.

Then, the negative SSIM loss, which can effectively measure the structural similarity between de-raining image and the clean one, is indicated as

$$L_{ssim} = 1 - \text{SSIM}(G(x), y) \tag{7}$$

where SSIM is the structural similarity index of two images, and a bigger score indicates a higher similarity.

Besides, the perceptual loss which can improve the visual quality of image processing results by minimizing the difference in high-level information [35] is indicated as

$$L_p = \sum_{i=3,8,15,25} E_x||V_i(G(x)) - V_i(y)||_2^2 \tag{8}$$

where $V_i(\cdot)$ denotes the output feature maps of the i-th layer of VGG16 [35] network pretrained on the ImageNet dataset.

Furthermore, the attention loss utilized both on the generator and discriminator can be respectively formulated as

$$L_{G\_att} = E_x||MAAM(x) - M(x)||_2^2 \tag{9}$$

$$L_{D\_Att} = E_x||D(x)_{mask} - M(x)||_2^2 \tag{10}$$

where MAAM(x) and $\text{D(x)}_{\text{mask}}$ are the attention map predicted by the generator and discriminator, respectively, and M(x) is a binary map of the rain streaks. In the binary map, zero represents the clean area and one represents the rainy area. In practice, the binary map is obtained by subtracting the ground-truth from the rainy image and then comparing the difference with a threshold.

The GAN loss for the relativistic generator is symmetric with Eq. (5) and is defined as

$$L_{G\_Ra} = -E_{x_r}[\log(1 - D_{Ra}(x_r, x_f))] - E_{x_f}[\log(D_{Ra}(x_f, x_r))] \tag{11}$$

Therefore, the total loss for the generator can be described as

$$L_G = \lambda_1 L_c + \lambda_2 L_{ssim} + \lambda_3 L_p + \lambda_4 L_{G\_Att} + \lambda_5 L_{G\_Ra} \tag{12}$$

where $\lambda_1 \sim \lambda_5$ are the coefficients to balance different loss terms and the total loss for the discriminator can be described as

$$L_D = L_{D\_Att} + L_{D\_Ra} \tag{13}$$

Equations (12) and (13) are minimized alternately through learning to optimize EAGAN to restore realistic scenes.

IV. EXPERIMENTS

In this section, the experimental settings and the quality measures utilized to evaluate the effectiveness of the proposed EAGAN are presented in detail, and then the deraining comparations between the EAGAN and the state-of-the-art methods on both synthetic and real-world datasets are executed. Finally, ablation studies are conducted to validate the improvements made by the main components in the proposed method.

A. EXPERIMENT SETTINGS

1) EXPERIMENT DETAILS

In this study, the entire network of EAGAN is implemented by the PyTorch framework on a Nvidia 1080Ti GPU. The Adam optimizer policy and the rectified linear unit (ReLU) are used as the loss optimization and the nonlinear operation, respectively. The learning rate is initialized as 5×10^{-4} and decayed by multiplying 0.2 when reaching the 30, 50 and 80 epochs. Besides, we set $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 1, \lambda_5 = 0.01$ as the coefficients to balance different loss terms in the experiments.

2) TRAINING DATASET

Considering that it is impossible to gain a large number of rainy images and the corresponding rain-free images in the real world, we apply three frequently used synthesized benchmark datasets, i.e., Rain100H [22], Rain100L [22], and Rain800 [25] to evaluate the proposed approach. Specifically, Rain100L consists of 200 image pairs for training and 100 image pairs for testing, the rainy images are synthetic with one type of rain streak direction in a single image, and the density of the rain is low. The rainy images in Rain100H are synthetic with five types of rain streak directions, the density of the rain is high and the background texture is badly damaged. Compared with Rain100L, Rain100H is more challenging, and its image pairs for training and testing are 1800 and 100, respectively. Rain800, whose rain-free images are randomly chosen from the UCID [36] and BSD-500 [37] and rainy images are synthesized by Photoshop, consists of 700 image pairs for training and 100 image pairs for testing. The background in Rain800 is more diverse and the density of rain is moderate compared with Rain100L and Rain100H.

Furthermore, Zhang *et al.* [25] and Yang *et al.* [22] provide abundant of real-world rainy images, whose rain streaks are varied in intensity and orientation. We use these datasets for objective evaluation to assess the generalization ability of the proposed EAGAN.

3) QUALITY MEASURES

Benefiting from that the ground-truth of synthesized datasets are available, we apply two frequently used reference metrics, i.e., structural similar index (SSIM) and peak signal to noise ratio (PSNR) to evaluate the deraining performance of different methods on synthetic images. However, the rain-free

TABLE 1. Average PSNR/SSIM of various methods on synthetic datasets. Bold values indicate the best results.

Dataset	Rain100H		Rain100L		Rain800	
Metrics	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
GMM	14.262	0.544	29.111	0.880	20.462	0.730
DSC	15.661	0.423	24.162	0.870	18.561	0.599
ReHEN	27.972	0.864	37.412	0.980	26.961	0.854
RESCAN	26.451	0.846	35.672	0.967	24.091	0.841
SPANet	26.762	0.836	35.473	0.967	25.960	0.866
UMRL	26.840	0.854	32.305	0.966	25.155	0.879
PReNet	27.790	0.873	37.111	0.976	26.295	0.857
JORDER-E	24.554	0.784	36.802	0.968	26.021	0.786
IAND	27.860	0.835	32.530	0.934	26.710	0.865
EAGAN	**29.980**	**0.902**	**37.476**	**0.981**	**26.967**	**0.893**

images of real-world datasets are not available, the deraining performance on these images can only be evaluated visually or subjectively. We compare the proposed EAGAN with two traditional optimization-based methods, i.e., DSC [11] and GMM [14], and several learning-based methods, i.e., ReHEN [24], RESCAN [23], SPANet [21], UMRL [28], PReNet [40], JORDER-E [41], and IAND [44].

B. EFFECTIVENESS ON SYNTHETIC DATASETS

Table 1 shows the quantitative comparison results of different methods on three synthetic datasets with different sizes and complexities. As can be seen, the deraining results of traditional methods are significantly lower than those of other learning-based methods in all metrics. Furthermore, the proposed EAGAN achieves a competitive performance on all datasets compared to other learning-based methods.

To visually illustrate the improvements obtained by EAGAN, deraining results of three images are presented in Fig. 5. As can be observed, DSC and GMM can hardly remove rain streaks and the background is still indistinguishable for all images. RESCAN, ReHEN, and PReNet can remove most rain streaks, but tend to blur image details and generate some unpleasant artifacts under difficult conditions, for example, the sky and cloud are heavily fused in the second deraining image. SPANet tends to remove image detail and remain rain streaks. UMRL can almost completely remove the rain streaks but tends to blur image details, for example, the branches in the second image and the windows in the third image are blurred heavily. Compared with other approaches, the proposed EAGAN can produce the best visual deraining performance by effectively removing the rain steaks while perfectly restoring the image structural details. Note that we choose the difficult sample images to test the models in Fig. 5, indicating that the proposed method performs well under difficult conditions.

In addition to the quantitative measures and visual effects, the proposed method is also efficient. Table 2 shows the total number of parameters involved in the evaluation and average running time on a rain image with a resolution of 320×640.

TABLE 2. Comparison of parameters number and average running time.

Methods	Time (seconds)	Parameters(millions)
GMM	423.511	X
DSC	146.761	X
ReHEN	0.443	0.298
RESCAN	0.522	**0.150**
SPANet	**0.145**	0.219
UMRL	0.162	0.984
PReNet	0.182	0.169
JORDER-E	0.410	4.169
IADN	0.167	0.980
EAGAN	0.165	3.901

For the two traditional optimization-based methods, GMM and DSC, they either need to learn the Gaussian mixture models or dictionary, which is very time consuming. In contrast, the learning-based methods are highly efficient benefiting from the fact that the pre-trained models are available. Due to the use of multi-channel, our EAGAN and JORDER-E have more parameters than the others, but they don't take much time, because other methods have some extra operations, for example, network reuse in ReHEN, RESCAN, SPANet, and PReNet, spatial recurrent neural network (RNN) in SPANet, and cycle spinning in UMRL.

C. EFFECTIVENESS ON REAL-WORLD DATASETS

To evaluate the generalization ability of the proposed EAGAN, the performance of different deraining methods on real-world rainy images are also compared. Fig. 6 represents the deraining results of all the mentioned methods on three real-world rainy images. As can be seen, DSC and GMM cannot effectively remove rain streaks and there are still residual rain streaks in the results. ReHEN, SPANet and PReNet have some improvements on rain streaks removal but tend to blur structural details, such as the gap between branches in the first image. RESCAN, UMRL, and JORDER-E tend to remain artifacts and cannot thoroughly remove rain streaks under difficult conditions. Compared with the above seven methods, the proposed EAGAN achieves the best performance on

FIGURE 5. Visual quality comparison of various methods on synthetic images. (a1-a3) Rainy images. (b1-b3) Results by GMM. (c1-c3) Results by DSC. (d1-d3) Results by ReHEN. (e1-e3) Results by RESCAN. (f1-f3) Results by SPANet. (g1-g3) Results by UMRL. (h1-h3) Results by PReNet. (i1-i3) Results by EAGAN. (j1-j3) Ground-truths (Please enlarge the images for better visualization).

rain streaks removal by restoring the most realistic texture information.

D. EFFECTIVENESS ON REAL RAINY VIDEOS

Furthermore, we evaluate the proposed EAGAN on real rainy videos and compare the performance with two state-of-the-art video deraining approaches, i.e., FastDerain [7] and MSCSC [8]. Since videos can be decomposed into successive frames, EAGAN processes videos in a frame-by-frame manner.

We downloaded a rainy video from the Internet and captured 1,000 successive frames for processing. The deraining results of frames 141, 153, 250, and 251 are shown in Fig. 7. Note that six frames adjacent to the 153th frame are out of order. As can be seen, for frame 141, FastDerain, MSCSC, and EAGAN can thoroughly remove the rain streaks, indicating that our EAGAN is effective even the temporal information is not available. However, the performance of both FastDerain and MSCSC degrade when dealing with the switched frames. Specifically, their deraining results of 153th

FIGURE 6. Visual quality comparison of various methods on real-world images. (a1-a3) Rainy images. (b1-b3) Results by GMM. (c1-c3) Results by DSC. (d1-d3) Results by ReHEN. (e1-e3) Results by RESCAN. (f1-f3) Results by SPANet. (g1-g3) Results by UMRL. (h1-h3) Results by PReNet. (i1-i3) Results by JORDER-E. (j1-j3) Results by EAGAN. (Please enlarge the images for better visualization).

frame contain residual rain streaks. Besides, we notice that when the video content was switched, i.e., the following frame described different content from the previous one, the performance of FastDerain and MSCSC are severely degraded. As shown in the fourth row of Fig. 7, FastDerain contains many residual rain streaks in the result while MSCSC generates unpleasant ghost of the previous frame. Compared with FastDerain and MSCSC, the proposed EAGAN is highly efficient and is not affected by the confused frame order or the changed content in adjacent frames.

E. ABLATION STUDIES

In this section, we conduct ablation studies to validate the main components of the proposed method. To this end, three variants of EAGAN are trained on Rain100H with the same strategy. The quantitative and qualitative results

FIGURE 7. Visual quality comparison on a real rainy video. The deraining results of frames 141, 153, 250 and 251 are from top to bottom. (a) Rainy images. (b) Results by FastDerain. (c) Results by MSCSC. (d) Results by EAGAN.

TABLE 3. Average PSNR/SSIM of the network with different settings. Bold values indicate the best results.

Variants	SSIM	PSNR
EAGAN-w/o-ATT	0.8879	29.0075
EAGAN-w/o-AGG	0.8962	29.1242
EAGAN-w-IN	0.7747	14.9664
EAGAN- w/o -SN	0.8904	29.8556
EAGAN-w/o-Re	0.8951	29.3631
EAGAN-w/o-D	0.8256	28.5134
EAGAN-w/o-NLNM	0.8635	29.3166
EAGAN-w/o-SE	0.8723	29.4486
EAGAN	**0.9023**	**29.9796**

of all the related variants are shown in Table 3 and Fig. 8, respectively.

1) EFFECTIVENESS ON ATTENTION MODULE AND AGGREGATION OPERATION

According to the subsections network architecture, the attention map can guide both the generator and discriminator to put conscious attention to rain regions. To validate the effectiveness of attention map, we remove the attention model from EAGAN. From Table 3, one can see that EAGAN gets a considerable margin both on the SSIM and PSNR metrics than EAGAN without attention module (EAGAN-w/o-ATT). In addition, visual quality comparisons in Fig. 8 also present that EAGAN can restore more image details. This demonstrates that the attention models enhance the deraining ability of the proposed EAGAN. Besides, we have studied a variant of EAGAN by removing the aggregation operation in the attention module (EAGAN-w/o-AGG), Table 3 shows that the aggregation operation can further improve the performance, which is consistent with the inference in Section B.

2) EFFECTIVENESS ON SPECTRAL NORMALIZATION

To validate the improvements obtained by spectral normalization, we have the variant EAGAN-w-IN, which introduces instance normalization for all convolution layers instead of spectral normalization.

From Table 3, one can see that EAGAN with instance normalization seriously degrades the SSIM and PSNR. Similarly, Fig. 8 presents the same result, in which EAGAN-w-IN produces a large number of unpleasant artifacts, the main reason is that instance normalization can easily make some impact on high-frequency feature information and cause some undesired artifacts. Therefore, we apply spectral normalization both in the generator and discriminator to ensure the training stability and restore realistic scenes, simultaneously. We also

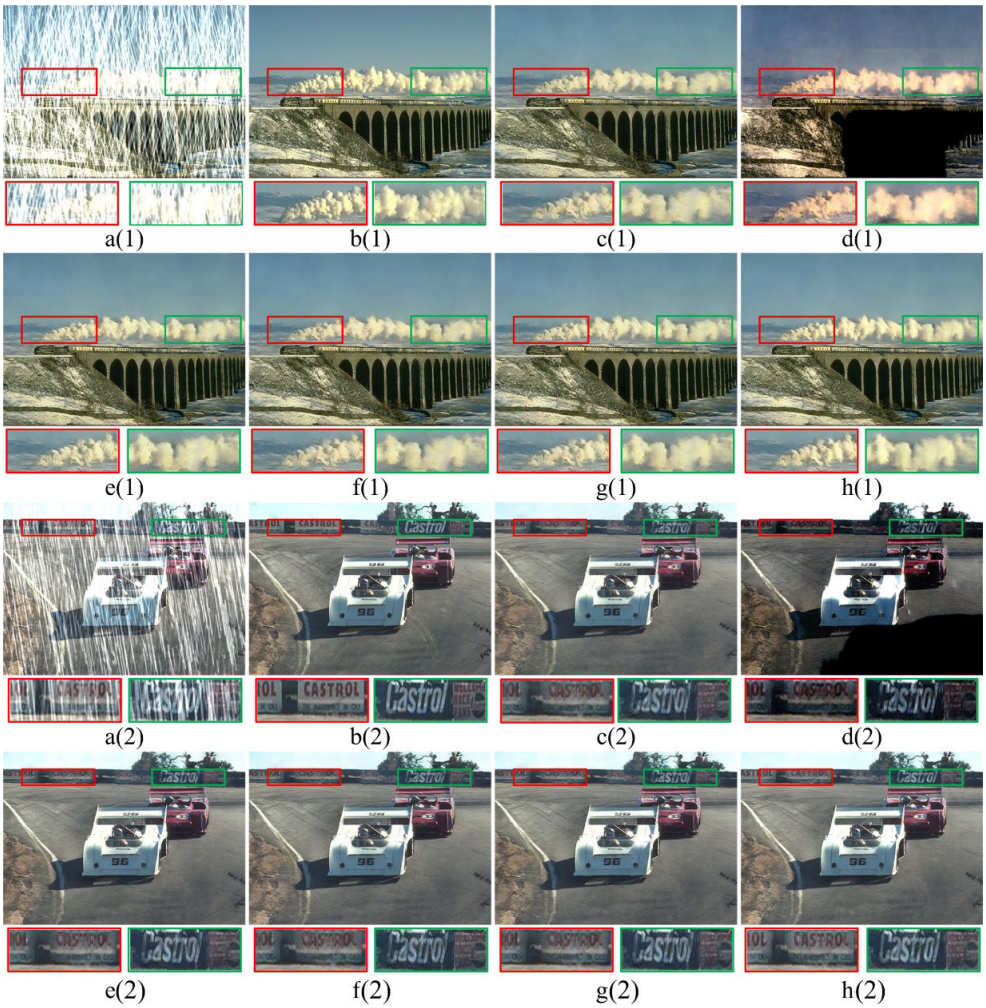


FIGURE 8. Visual quality comparison of the network with different network architecture settings. (a) Rainy images. (b) Ground-truths. (c) Results by EAGAN-w/o-Att. (d) Results by EAGAN -w-IN. (e) Results by EAGAN -w/o-Re. (f) Results by EAGAN -w/o-NLNM. (g) Results by EAGAN -w/o-SE. (h) Results by EAGAN.

test a variant of EAGAN by removing spectral normalization, i.e., EAGAN- w/o -SN in Table 3, one can see that spectral normalization can further improve the performance, besides, spectral normalization can enhance the training stability [39].

3) EFFECTIVENESS ON RELATIVISTIC GAN AND DISCRIMINATOR

The effectiveness of relativistic GAN in EAGAN is studied. To this end, the deraining comparisons of traditional GAN (EAGAN-w/o-Re) with EAGAN are shown in Table 2 and Fig. 8, respectively. From Table 3, one can see that EAGAN can further improve the performance both on PSNR and SSIM metrics. Besides, Fig. 8 illustrates that EAGAN can restore more image details than EAGAN-w/o-Re, indicating the benefit of relativistic GAN in the proposed method. Furthermore, the discriminator is removed from the network, leaving only the generator to remove rain streaks. From Table 3, one can see that the deraining performance without discriminator (EAGAN-w/o-D) is severely degraded, indicating that the discriminator can promote the generator greatly.

4) EFFECTIVENESS ON NLNM AND SE MOUDLE

Finally, the effectiveness of NLNM and SE module in generator is studied. From Table 3, one can see that the deraining performance of EAGAN-w/o-NLNM and EAGAN-w/o-SE suffer some degradation compared to EAGAN, Fig. 8 shows the same results, indicating that both NLNM and SE modules can improve the representation ability of the deraining network.

V. CONCLUSION

In this work, we have proposed an enhanced attention generative adversarial network named EAGAN to remove rain streaks from a single image. In EAGAN, a multiscale aggregation attention module (MAAM) with various convolution kernel sizes is developed to locate the rain regions, and then non-local neural networks are embedded to cap-

ture long-range spatial dependencies and structural information. Furthermore, symmetrical autoencoder with long-range skip-connections and SE modules enhance the representation power of the network. Besides the improvement in the generator, we adopt the relativistic discriminator to further improve the performance. Meanwhile, spectral normalization is utilized in both the generator and discriminator to enhance the training stability and produce realistic deraining scenes. Extensive experiments on both synthetic and real-world datasets have shown the effectiveness of the proposed EAGAN in terms of quantitative and qualitative comparisons.

REFERENCES

[1] S. Liu, D. Huang, and Y. Wang, "Receptive field block net for accurate and fast object detection," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 304–419, doi: 10.1007/978-3-030-01252-6_24.

[2] F. S. Saleh, M. S. Aliakbarian, M. Salzmann, L. Petersson, and J. M. Alvarez, "Effective use of synthetic data for urban scene semantic segmentation," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 86–103, doi: 10.1007/978-3-030-01216-8_6.

[3] Y. Miao, H. Hong, and H. Kim, "Size and angle filter based rain removal in video for outdoor surveillance systems," in *Proc. 8th Asian Control Conf. (ASCC)*, May 2011, pp. 1300–1304.

[4] K. Garg and S. K. Nayar, "Detection and removal of rain from videos," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2004, p. 528, doi: 10.1109/CVPR.2004.1315077.

[5] K. Garg and S. K. Nayar, "Vision and rain," *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 3–27, Jul. 2007.

[6] J.-H. Kim, J.-Y. Sim, and C.-S. Kim, "Video deraining and desnowing using temporal correlation and low-rank matrix completion," *IEEE Trans. Image Process.*, vol. 24, no. 9, pp. 2658–2670, Sep. 2015, doi: 10.1109/TIP.2015.2428933.

[7] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, and Y. Wang, "A novel tensor-based video rain streaks removal approach via utilizing discriminatively intrinsic priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2818–2827, doi: 10.1109/CVPR.2017.301.

[8] M. Li, Q. Xie, Q. Zhao, W. Wei, S. Gu, J. Tao, and D. Meng, "Video rain streak removal by multiscale convolutional sparse coding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 6644–6653, doi: 10.1109/CVPR.2018.00695.

[9] M. Li, X. Cao, Q. Zhao, L. Zhang, and D. Meng, "Online rain/snow removal from surveillance videos," *IEEE Trans. Image Process.*, vol. 30, pp. 2029–2044, 2021, doi: 10.1109/TIP.2021.3050313.

[10] J.-H. Kim, C. Lee, J.-Y. Sim, and C.-S. Kim, "Single-image deraining using an adaptive nonlocal means filter," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 914–917, doi: 10.1109/ICIP.2013.6738189.

[11] Y. Luo, Y. Xu, and H. Ji, "Removing rain from a single image via discriminative sparse coding," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3397–3405, doi: 10.1109/ICCV.2015.388.

[12] Y.-L. Chen and C.-T. Hsu, "A generalized low-rank appearance model for spatio-temporally correlated rain streaks," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1968–1975, doi: 10.1109/ICCV.2013.247.

[13] S. Gu, D. Meng, W. Zuo, and L. Zhang, "Joint convolutional analysis and synthesis sparse representation for single image layer separation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1717–1725, doi: 10.1109/ICCV.2017.189.

[14] Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown, "Rain streak removal using layer priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2736–2744, doi: 10.1109/CVPR.2016.299.

[15] X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley, "Removing rain from single images via a deep detail network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1715–1723, doi: 10.1109/CVPR.2017.186.

[16] D. Zhang, J. He, and Y. Zhao, "Learning to perform joint image super-resolution and rain removal via a single-convolutional neural network," *J. Electron. Imag.*, vol. 27, no. 6, Dec. 2018, Art. no. 063024, doi: 10.1117/1.JEI.27.6.063024.

[17] H. Xia, R. Zhuge, H. Li, S. Song, F. Jiang, and M. Xu, "Single image rain removal via a simplified residual dense network," *IEEE Access*, vol. 6, pp. 66522–66535, 2018, doi: 10.1109/ACCESS.2018.2879330.

[18] W. Yang, J. Liu, S. Yang, and Z. Guo, "Scale-free single image deraining via visibility-enhanced recurrent wavelet learning," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2948–2961, Jun. 2019, doi: 10.1109/TIP.2019.2892685.

[19] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the skies: A deep network architecture for single-image rain removal," *IEEE Trans. Image Process.*, vol. 26, no. 6, pp. 2944–2956, Jun. 2017, doi: 10.1109/TIP.2017.2691802.

[20] Y. Chang, L. Yan, and S. Zhong, "Transformed low-rank model for line pattern noise removal," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 1735–1743, doi: 10.1109/ICCV.2017.191.

[21] T. Wang, X. Yang, K. Xu, S. Chen, Q. Zhang, and R. W. H. Lau, "Spatial attentive single-image deraining with a high quality real rain dataset," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 12262–12271, doi: 10.1109/CVPR.2019.01255.

[22] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "Deep joint rain detection and removal from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 1685–1694, doi: 10.1109/CVPR.2017.183.

[23] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, "Recurrent squeeze-and-excitation context aggregation net for single image deraining," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2018, pp. 262–277-1694, doi: 10.1007/978-3-030-01234-2_16.

[24] Y. Yang and H. Lu, "Single image deraining via recurrent hierarchy enhancement network," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 1814–1822, doi: 10.1145/3343031.3351149.

[25] H. Zhang, V. Sindagi, and V. M. Patel, "Image de-raining using a conditional generative adversarial network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 11, pp. 3943–3956, Nov. 2020, doi: 10.1109/TCSVT.2019.2920407.

[26] R. Qian, R. T. Tan, W. Yang, J. Su, and J. Liu, "Attentive generative adversarial network for raindrop removal from a single image," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 2482–2491, doi: 10.1109/CVPR.2018.00263.

[27] G. Chai, Z. Wang, G. Guo, Y. Chen, Y. Jin, W. Wang, and X. Zhao, "Recursive modified dense network for single-image deraining," *J. Electron. Imag.*, vol. 29, no. 3, May 2020, Art. no. 033006, doi: 10.1117/1.JEI.29.3.033006.

[28] R. Yasarla and V. M. Patel, "Uncertainty guided multi-scale residual learning-using a cycle spinning CNN for single image de-raining," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8397–8406, doi: 10.1109/CVPR.2019.00860.

[29] L. Zhu, C.-W. Fu, D. Lischinski, and P.-A. Heng, "Joint bi-layer optimization for single-image rain streak removal," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2545–2553, doi: 10.1109/iccv.2017.276.

[30] X. Wang, R. Girshick, A. Gupta, and K. He, "Non-local neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7794–7803, doi: 10.1109/CVPR.2018.00813.

[31] D. Liu, B. Wen, Y. Fan, C. C. Loy, and T. S. Huang, "Non-local recurrent network for image restoration," 2018, *arXiv:1806.02919*. [Online]. Available: http://arxiv.org/abs/1806.02919

[32] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141, doi: 10.1109/CVPR.2018.00745.

[33] A. Jolicoeur-Martineau, "The relativistic discriminator: A key element missing from standard GAN," 2018, *arXiv:1807.00734*. [Online]. Available: http://arxiv.org/abs/1807.00734

[34] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: 10.1109/TIP.2003.819861.

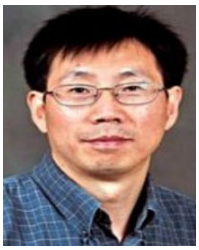
[35] J. Johnson, A. Alahi, and F. F. Li, "Perceptual losses for real-time style transfer and super-resolution," in *Computer Vision—ECCV* (Lecture Notes in Computer Science), vol. 9906, B. Leibe, J. Matas, N. Sebe, and M. Welling, Eds. 2016, pp. 694–711), doi: 10.1007/978-3-319-46475-6_43.

[36] G. Schaefer and M. Stich, "UCID: An uncompressed color image database," *J. Electron. Imag.*, vol. 5307, pp. 472–480, Dec. 2003, doi: 10.1117/12.525375.

[37] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011, doi: 10.1109/TPAMI.2010.161.
[38] H. Zhang and V. M. Patel, "Density-aware single image de-raining using a multi-stream dense network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 695–704, doi: 10.1109/CVPR.2018.00079.
[39] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks," in *Proc. Int. Conf. Learn. Represent.*, 2018, pp. 1–29. [Online]. Available: https://arxiv.org/abs/1802.05957
[40] D. Ren, W. Zuo, Q. Hu, P. Zhu, and D. Meng, "Progressive image deraining networks: A better and simpler baseline," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 3932–3941, doi: 10.1109/CVPR.2019.00406.
[41] W. Yang, R. T. Tan, J. Feng, Z. Guo, S. Yan, and J. Liu, "Joint rain detection and removal from a single image with contextualized deep networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 6, pp. 1377–1393, Jun. 2020, doi: 10.1109/TPAMI.2019.2895793.
[42] W. Yang, R. T. Tan, S. Wang, Y. Fang, and J. Liu, "Single image deraining: From model-based to data-driven and beyond," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, May 19, 2020, doi: 10.1109/TPAMI.2020.2995190.
[43] Y. Du, J. Xu, Q. Qiu, X. Zhen, and L. Zhang, "Variational image deraining," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 2395–2404, doi: 10.1109/WACV45572.2020.9093393.
[44] K. Jiang, Z. Wang, P. Yi, C. Chen, Z. Han, T. Lu, B. Huang, and J. Jiang, "Decomposition makes better rain removal: An improved attention-guided deraining network," *IEEE Trans. Circuits Syst. Video Technol.*, early access, Dec. 15, 2020, doi: 10.1109/TCSVT.2020.3044887.
[45] K. Jiang, Z. Wang, P. Yi, C. Chen, B. Huang, Y. Luo, J. Ma, and J. Jiang, "Multi-scale progressive fusion network for single image deraining," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 8343–8352, doi: 10.1109/CVPR42600.2020.00837.
[46] G. Chai, Z. Wang, G. Guo, Y. Chen, Y. Jin, W. Wang, and X. Zhao, "Recurrent attention dense network for single image de-raining," *IEEE Access*, vol. 8, pp. 111278–111288, 2020, doi: 10.1109/ACCESS.2020.3003126.

GUOQIANG CHAI received the bachelor's and master's degrees from the University of Electronic Science and Technology of China, in 2011 and 2014, respectively. He is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, North University of China. His main research interests include image processing and machine learning.

ZHAOBA WANG received the Ph.D. degree in instruments science and technology from the Nanjing University of Science and Technology, in 2002. He is currently a Professor with the School of Information and Communication Engineering, North University of China. His main research interests include information processing and reconstruction, computer vision, and image processing.



GUODONG GUO (Senior Member, IEEE) received the Ph.D. degree in pattern recognition and intelligent control from the Chinese Academy of Sciences and the Ph.D. degree in computer science from the University of Wisconsin–Madison. His main research interests include machine learning, computer vision, multimodal image classification, video surveillance and scene understanding, and human behavior analysis.

YOUXING CHEN received the Ph.D. degree from the North University of China, in 2010. He is currently a Professor with the School of Information and Communication Engineering, North University of China. His research interests include the areas of image processing, signal processing, and non-destructive testing.

YONG JIN received the Ph.D. degree from the North University of China, in 2013. He is currently a Professor with the School of Information and Communication Engineering, North University of China. His research interests include the areas of image processing, online inspections, and big data analytics.

DAWEI WANG received the Ph.D. degree from the North University of China, in 2020. He is currently a Teacher with the School of Physics and Information Engineering, Shanxi Normal University. His main research interests include image processing, signal processing, and ultrasonic signal processing.

BIN LU received the B.S. and Ph.D. degrees from the School of Microelectronics, Xidian University, Xi' an, China, in 2013 and 2019, respectively. He is currently with the School of Physics and Information Engineering, Shanxi Normal University, Linfen, China. His current research interest includes modeling, design, and fabrication of the tunneling field-effect-transistor (TFET).

SHILEI REN is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering, North University of China. His main research interests include image processing, signal processing, and medical signal processing.