Journal of Computers Vol. 34 No. 2, April 2023, pp. 129-140
doi: 10.53106/199115992023043402010

Multi-scale Fusion Residual Network For Single Image Rain Removal

Jia-Chen He[1], Ming-Jian Fu[2], Li-Qun Lin[1*]

[1] College of Physics and Information Engineering, Fuzhou University,
Fuzhou 350000, China
061900119@fzu.edu.cn, Lin_liqun@fzu.edu.cn

[2] College of Computer and Data Science, Fuzhou University,
Fuzhou 350000, China
sinceway@fzu.edu.cn

Received 13 May 2022; Revised 6 September 2022; Accepted 28 October 2022

Abstract. Deep learning has been widely used in single image rain removal and demonstrated favorable universality. However, it is still challenging to remove rain streaks, especially in the nightscape rain map which exists heavy rain and rain streak accumulation. To solve this problem, a single image nightscape rain removal algorithm based on Multi-scale Fusion Residual Network is proposed in this paper. Firstly, based on the motion blur model, evenly distributed rain streaks are generated and the dataset is reconstructed to solve the lack of nightscape rain map datasets. Secondly, according to the characteristics of the night rain map, multi-scale residual blocks are drawn on to reuse and propagate the feature, so as to exploit the rain streaks details representation. Meanwhile, the linear sequential connection structure of multi-scale residual blocks is changed to a u-shaped codec structure, which tackles the problem that features cannot be extracted effectively due to insufficient scale. Finally, the features of different scales are combined with the global self-attention mechanism to get different rain streak components, then a cleaner restored image is obtained. The quantitative and qualitative results show that, compared to the existing algorithms, the proposed algorithm can effectively remove rain streaks while retaining detailed information and ensuring the integrity of image information.

Keywords: deep learning, nightscape rain removal, Multi-scale Fusion Residual Network, global self-attention mechanism

1 Introduction

With the rapid development of computer vision, many computer vision technologies have been applied to deal with practical problems in daily life. Such as pedestrian detection [1], object tracking [2], semantic segmentation [3], etc. Since the video or image is captured outdoors, the quality of the image and the subsequent image processing effect can be decreased for a large number of reasons. Among them, rain is one of the common outdoor factors affecting image quality. With the universality of autonomous driving and more outdoor applications based on computer vision, rain removal becomes an integral part of low-level computer tasks once again.

In recent years, researchers have proposed many effective algorithms for the problem of removing rain from a single image. Yang et al. [4] divided them into two categories: model-based methods and data-driven methods. Between them, model-based methods can be further classified into filtering-based methods and prior-based methods. Using a single de-rain as a signal filtering task is a filter-based approach that uses edge retention and physical property filtering to recover images with rain [5-6]. However, the solution equation is considered as an optimization problem in the prior-based methods. Besides, the prior-based method applies traditional image processing methods such as discriminative sparse [6-7], Gaussian mixture model [8], and low-rank representation [7]. Unlike model-based methods, data-driven methods view rain removal as a process of learning a nonlinear function and map rain streaks into the background image by finding suitable parameters [4].

Convolutional Neural Networks (CNN) have shown good performance in image rain removal with promising successes in recent years [9-11]. [12] considered the location information of rain streaks in the image by learning the rain content at different scales and using them to estimate the final de-rained output. [13] proposed a scale aggregation module to learn features at different scales. At the same time, feature aggregation can adapt to

* Corresponding Author

each channel and has a superior de-noising effect because of the introduction of the attention module. [14] used the proximal gradient descent technique to design an iterative algorithm containing only simple operators and proposed a deep neural network with fully interpretable. A VNet network structure containing a spatial pyramid is described to predict water vapor propagation based on rain patterns and ANet's encoder is utilized to estimate atmospheric light [15]. [16] designed an optimization model-driven deep CNN. The unsupervised loss function of the optimization model is enforced on the proposed network for better generalization. A novel dual heterogeneous complementary networks are proposed to explore both detailed structure and semantic contexts [17]. Numerous studies have demonstrated that neural networks can bring significant performance improvement in image rain removal.

Most of the above studies were conducted for daytime rain maps. However, the nightscape rain map differs from the daytime rain map in two ways: on the one hand, the rain removal network based on the daytime rain map dataset will remove rain streaks excessively in the nightscape rain map, and on the other hand, the lamps in the nightscape can easily be removed by mistake because of the similarity of light and rain streaks at night. Considering the unique features of nightscape rain maps, there is still a lack of thorough research on rain removal in nightscape maps, and there is no public nightscape rain map dataset. In this paper, a rain removal network for nightscape rain maps is built by using multi-scale fusion residual blocks.

The major contributions of our work are summarized as follows:

1) A large dataset of nightscape rain maps is developed with digital image processing methods, which extends existing datasets and handles the lack of nightscape rain map datasets. Besides, it does not require a lot of effort.

2) A deep learning approach with a Multi-scale Fusion Residual Network (MFR-Net) for tiny rain streaks feature extraction is proposed. It results in an end-to-end model that jointly learns multi-scale rain streak features to remove rain streaks efficiently.

3) Superior performance of our method is achieved against popular rain removal algorithms. In addition, our algorithm proves the soundness of each module in ablation experiments, which shows that our algorithm has good generalization and robustness.

The rest of this paper is organized as follows. In Section 1, we propose a rain removal algorithm in the nightscape rain map. In Section 2, a method of synthesizing rain using a digital image processing algorithm is described. In Section 3, we introduce Multi-scale Attentive Residual Dense Network (MARD-Net) [18], Global Self-attention (GSA) module, and explain MFR-Net in detail. In Section 4, the relevant experimental results are given. Finally, this paper is concluded in Section 5.

2 Construction of the Dataset

In image de-raining models, a rainy image is considered to be a linear superposition of a clean (no rain) background image and a rain component. Equation 1 is expressed as:

$$I = B + R, \tag{1}$$

where B, R and I represents the background image, the rain component and the synthesized image, respectively.

In the real world, a large number of rain streaks are accumulated from various sizes of rain streaks. Rain streaks of the same size will show in different sizes depending on the photographer's viewpoint. When the rain streaks are densely accumulated, individual rain streaks will overlap and intersect. This accumulated rain streak has a visual effect, which is similar to haze or fog for distant scenes, leading to atmospheric occlusion and blurring effects. Therefore, the real rain streaks are approximated through random generation, superposition, rotation and blur. Besides, while ensuring the diversity of nightscape rain dataset, a large number of nightscape rain maps in batches are generated.

2.1 Generation of Rain Streaks

In the process of synthesizing the dataset, the key is how to build the R layer. Since rain presents approximately linear information in free fall, this paper considers the R layer as point-like noise that increases motion blur. A random number is generated between 0-255 as the generated pixel noise, and it is multiplied by the scale factor to control the size of the generated noise. In order to make the noise raindrops more realistic, the noise informa-

tion of the generated pixel is convolved with a blur matrix, so that the pixel noise has a preliminary motion blur effect.

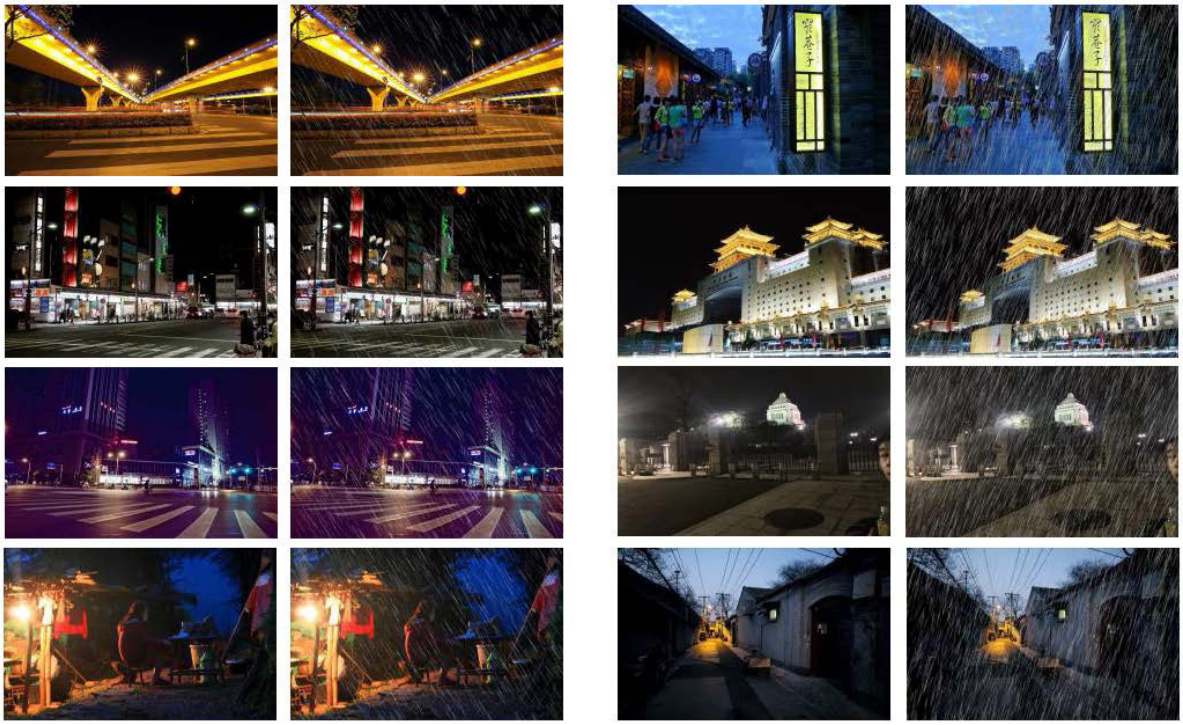
On the basis of obtaining pixel point noise, point noise can be transformed into line noise through affine transformation, thereby generating rain streak. In one-dimensional space, for any vector x, the translation amount is b, and the affine transformation of the zoom factor A is as follows:

$$y = Ax + b. \tag{2}$$

Generalize Equation 2 to a two-dimensional image, assuming the origin is O, $f(\mathrm{O}) \in B$, B denoted the original image, then for any vector x it is represented as:

$$f(\mathrm{O} + \mathrm{x}) \rightarrow \varphi(B + \mathrm{x}). \tag{3}$$

After the noise map is obtained, a diagonal matrix d is generated, and the noise map is affinely transformed to the diagonal matrix d at this time, the raindrops are transformed into rain streaks. Next, use Gaussian blurring the rain pattern to make the rain pattern possess a "width", so that the complete rain component R is obtained.



Clean image Rainy image Clean image Rainy image

(a) Training set (b) Testing set

Fig. 1. Synthetic dateset

2.2 Construction of Synthetic Datasets

According to Equation 1, after getting the rain component R, the final rain-containing image I can be obtained only by selecting clean background picture B for channel superposition. In this method, the rain information is superimposed on each channel. The part representing black in each channel of the rain component R will give different weights to the value on the channel at the same position in the background image B. The weights are adjusted according to the distance of the simulated raindrop to the lens. This method based on different weights

makes the final image appear as clear or blurry rain streaks, increasing the spatial information of rain. The final image produced by this method does not change the brightness information of daytime and night images. In this paper, the channel superposition method is used to reconstruct the night and day rain datasets required for the experimental part of this paper.

For the nightscape rain map dataset, since it is difficult to obtain rain image datasets by cameras in the real environment, a widely acceptable method for constructing datasets is to apply synthetic rain to construct datasets. The nightscape rain map dataset Rain_dark100L constructed in this paper adopts 200 city night background images as the base images of the training set. The rain streaks are set to be lightweight, including the angle range from -30° to 30°, and various rain streak lengths; the test set utilizes rain streaks addition twice, including two rain streaks from -30° to 30°. Multiple experiments are proved that when two kinds of rain streaks information are added to the same image, the network learning effect is the best. This paper adds the rain streak information twice on the test set. After two additions, two different kinds of rain streak information will appear in the images in the test set and are used to test the learning effect of each round of rain streak information of the network. Rain_dark100H adopts the same addition mode as Rain_dark100L, but increases the density of rain streaks.

Scholars are now widely applying the Rain100L [19], Rain100H [19], and Rain14000 [20] datasets. Among them, Rain100L contains 1800 pairs of rainy/clean images as a training set, and 200 pairs of rainy/clean images as a test set. Rain100H and Rain100L have the same number of images in the training set and test set, and the background images are also the same, but the densities of rain streaks are different. Rain14000 contains a training set of 12,600 pairs of images with/without rain, and a test set of 1,400 pairs of images with/without rain.

In order to verify the effectiveness of the algorithm in this paper, while constructing the nightscape dataset, the method was used to reconstruct the dataset Rain100L and Rain100H renamed as Rain2000L and Rain2000H. Fig. 1 presents part of synthetic rain datasets.

3 The Proposed Multi-scale Fusion Residual Network

The performance of existing rain removal algorithms is limited due to excessive smoothing effect, poor generalization ability, and differences in rainfall intensity in spatial location and color channel. This paper proposes an end-to-end based multi-scale fusion residual dense network. The network model uses a completely symmetric U-shaped structure to enable feature fusion and further integrates the Multi-scale Fusion Residual Blocks into the U-shaped network to improve the capabilities of rain feature extraction and representation. Also, the GSA module is introduced before output, which can better extract the information of the tiny rain streak, so as to achieve an overall improvement in the rain removal performance. In the following subsections, Multi-scale attention residual block, GSA module and Multi-scale fusion residual network will be discussed in detail. The network architecture is illustrated in Fig. 2.

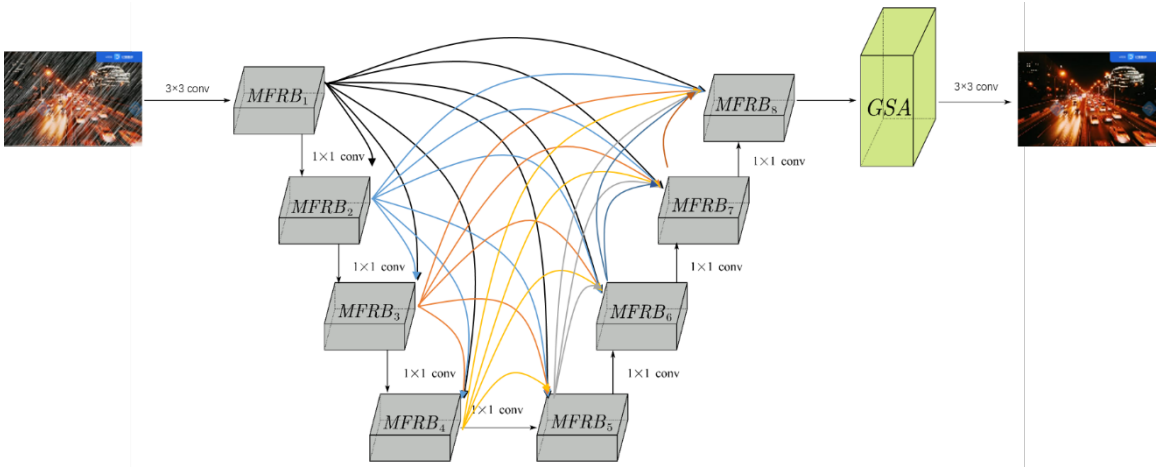


Fig. 2. The structure of MFR-Net

3.1 Multi-scale Attention Residual Block

The multi-scale residual network is composed of basic Multi-Scale Attentive Residual Blocks (MARB). This module can take rain images of different scenes as input, and it can well represent rain streak features after MARB processing. The input features of MARB are concatenated after passing through a 3×3 and 5×5 convolutional layers so that to complete multi-scale information fusion. In order to further improve the representation capability of the network, MARB also introduces inter-layer multi-scale information fusion to fuse multi-scale information with features of different scales. This structure enables MARB to learn the main features of images through different scales and features, ensuring that the input information of the network can be propagated through all parameter layers. In order to speed up the training process of the network, this paper also introduces a global skip connection between different MARBs, and updates the parameters by calculating the back-propagation gradient. This skip connection method can directly propagate the lossless information in the whole network, and realize the effective extraction of the final image.

Since the distributions of rain density are varied greatly among different color channels, Bright Channel Prior (BCP) is also adopted in MARD-Net to obtain the feature weighting information of different color channels [21]. The BCP-based channel attention scheme [22] can assist the network to maintain the pixel brightness of the extracted image, helping to extract crucial features of the image. Moreover, due to the inhomogeneous distribution of rain streaks, there may be differences in different spatial locations. In this paper, multi-scale information fusion is achieved through convolutional layers of 1×1 and 3×3, and chan-nel-level and spatial-level attention modules are introduced to notice the rainfall areas.

3.2 Attentional Mechanism

Attention Mechanism, as an overwhelmingly paramount part of neural network research, has invariably been concerned by scholars. The transformer [23-25] was recently introduced in the field of Natural Language Processing (NLP), its performance is extremely excellent. The traditional CNN and RNN architectures are no longer applied in transformer. The entire network structure is completely composed of the Attention mechanism, which has better parallelism and conforms to the existing GPU framework. Thus, the Attention mechanism replaces the convolutional layer in this paper.

In comparison to the Attention mechanism, the Self-attention mechanism, which is the core of the Transformer, reduces the dependence on external information and is good at capturing the internal correlation of data or features. Self-attention mechanism is widely used in one-dimensional language processing in the field of NLP, and is also extended to the two-dimensional image, such as VIT [24]. Self-attention mechanism adopts the calculation method of scaled dot-product, as expressed in Equation 4:

$$f(\mathbf{Q},\mathbf{K},\mathbf{V})=\text{softmax}(\frac{\mathbf{Q}\mathbf{K}^{\mathrm{T}}}{\sqrt{\mathrm{d_k}}})\mathbf{V}, \tag{4}$$

where $\boldsymbol{Q}$, $\boldsymbol{K}$ and $\boldsymbol{V}$ are the query vector, the key vector, and the value vector respectively. These three vectors are created by multiplying word embedding with three weight matrices.

Global self-attention mechanism (GSA) [26] works similarly to Self-attention mechanism. GSA also exploits three vectors of key, query and value to generate new features, and is focus on the content and spatial position of pixels through the content attention layer and the position attention layer. Equation 5 gives the calculation method of GSA:

$$F = Q\,(\,\rho(\mathbf{K}^{T})\,\mathbf{V}), \tag{5}$$

where $\mathbf{K}^{\mathrm{T}}$ is the transposed matrix of $\mathbf{K}$, ρ denotes the softmax normalization operation for each row, and the Q function means that the pixels in $\mathbf{V}$ utilize the $\rho(\mathbf{K}^{T})$ weight.

In this paper, using GSA to replace Self-attention mechanism can resolve the problems of secondary computation and storage complexity. Furthermore, GSA can better extract rain features based on the ability to interact with remote pixels.

3.3 Multi-scale Fusion Residual Network Architecture

Although the MARD network makes good use of 3×3 and 5×5 convolution kernels and 1×1 and 3×3 convolution kernels to achieve multi-scale information fusion. Nevertheless, the features are only limited to the stacking of 3×3 and 5×5 convolutional layers, and there is still a problem of insufficient scale. Inspired by Pyramid Attention Networks (PANet) [27] and U-Net [28], this paper proposes a Multi-scale Fusion Residual Network (MFR-Net), a structure diagram of the MFR-Net is given in Fig. 2. MFR-Net changes the linear sequential connection structure in the original MARD-Net to the U-shaped Encoder-Decoder structure. When going down to the Encoder, the original 3×3 and 5×5 convolutional layers are changed to double the number of kernels every time they pass through one layer. Denote the input feature of Multi-scale Fusion Residual Block (MFRB) as $\mathbf{F}_0$, after different convolution kernels (respectively 1×1 and 3×3), the output results are presented in equations 6 and 7:

$$\mathbf{F}_1^{1\times1} = f_{1\times1}(\mathbf{F}_0; \mu_0^{1\times1}), \tag{6}$$

$$\mathbf{F}_1^{3\times3} = f_{3\times3}(\mathbf{F}_0; \mu_0^{3\times3}), \tag{7}$$

among them, $\mathbf{F}_1^{n\times n}$ indicates that the output size of the convolution is n × n , $f_{n\times n}(\cdot)$ indicates that the size of the convolution is n × n , and $\mu_0^{n\times n}$ indicates that the size of the hyperparameter convolution is n × n . The activation functions of these convolutional layers mentioned above generally utilize Leaky-ReLU, α = 0.2.

The outputs of the above equations are further passed through concat, which are given by equations 8:

$$\mathbf{F}_1^{'} = \quad (\mathbf{F}_1^{1\times1} + \mathbf{F}_1^{3\times3}; \mu_1) + \mathbf{F}_0, \tag{8}$$

where C is for performing a concat operation, and μ_1 is the hyperparameter. At this point, the output is corrected to obtain equation 9:

$$\mathbf{F}_{\text{out}} = sa(ca(C(\mathbf{F}_1^{1\times1} + \mathbf{F}_1^{3\times3}; \mu_1); \vartheta_0); \vartheta_1) + \mathbf{F}_0, \tag{9}$$

where $\mathbf{F}_{\text{out}}$ denotes the output of MFRB, $sa(\cdot)$ and $ca(\cdot)$ are the spatial attention mechanism and the channel attention mechanism, respectively, and $\{\mu_1\ \vartheta_0;\ \vartheta_1\}$ are the hyperparameters of the MFRB output. Using this network structure allows the network to better explore and reorganize features at different scales. Fig. 3 presents a schematic diagram of the internal module structure of MFRB.

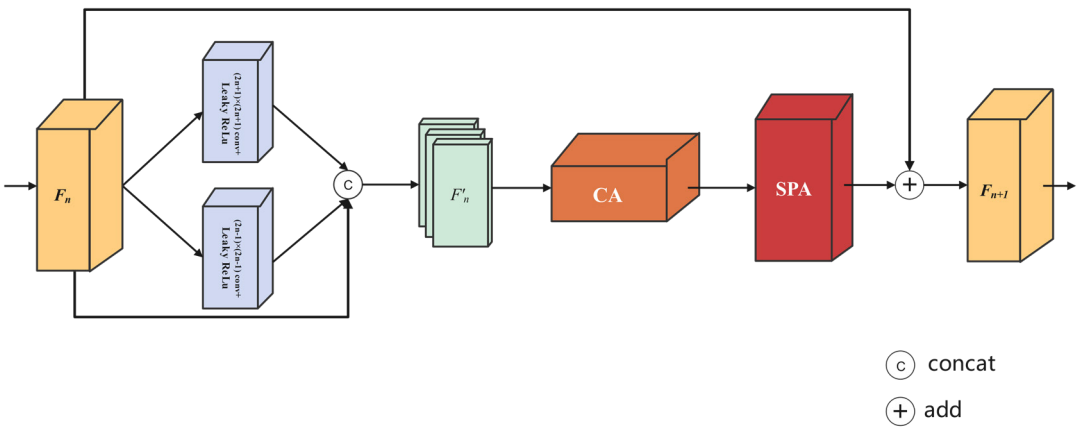


Fig. 3. MFRB internal structure

3.4 GSA Enhancement Module

GSA is introduced in this paper because of its superiority in quadratic computation and storage complexity and its ability to model long-range pixel interactions in the entire network. Through experiments, it is found that if GSA is introduced before and after each convolutional layer or directly replaces all convolutional layers with GSA, some no-rain features will be over-enhanced when the image feature is output, resulting in the overall dark or blurred image. However, when it is added to the last layer before the output, the network has learned the main features of rain. Therefore, connecting GSA is able to strengthen the feature, so that the original difficulties to recognize tiny rain streaks are removed well, more conducive to the learning of rain features.

The final improved network structure is shown in Fig. 2. When the picture is input, adopt 3 × 3 convolution for channel transformation, and transform the input 3 channels (RGB) into 32 channels. A skip connection is performed after each MFRB, so that the features between layers can be fully fused by the residual network. After each skip layer connection, a 1×1 convolution channel transformation is added to change the 64 channels obtained by the skip layer connection back to 32 channels. After 8 MFRBs, access a GSA module for feature enhancement, and then go through a 3×3 convolution to change back to 3 channels to get the final output.

4 Experiment and Analysis

In order to verify the effectiveness of the algorithm in this paper, the nightscape datasets selected in this paper are Rain_dark100L, Rain_dark100H and Rain_dark9000. Among them, Rain_dark100L and Rain_dark100H include 1800 pairs of rainy/clean training set images and 200 pairs of rainy/clean test set images. Rain_dark9000 includes 8400 pairs of rainy/clean training set images and 600 pairs of rainy/clean test set images. The reconstructed daytime rain datasets Rain2000L and Rain2000H both contain 1800 pairs of rainy/clean images as training set and 200 pairs of rainy/clean images as test set.

4.1 Experimental Environment and Parameter Setting

All the experimental results in this paper are completed on the same server running Linux (Ub-untu18.04). The CPU is Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40GHz, the GPU is Nvidia RTX-2080Ti, and the training platform is built on the deep learning framework Pytorch1.7.1 based on IDLE (Python 3.7 64-bit). In order to ensure the consistency of the experiment, this paper applies our proposed datasets Rain_dark100L, Rain_dark100H and Rain_dark9000 for model training, the learning rate is set to 0.01, and 200 batches are used, and the trained model is tested for the real rain maps.

4.2 Validation of Dataset Reliability

In order to verify the effectiveness of the proposed dataset construction method, this paper utilizes the same background image to construct training set and test set. And tested in MARD-Net, GridDerainNet [29], SPANet [30] and MSPFN [31], using the same real picture to test the actual rain removal effect. Finally, the experimental results in Table 1, Table 2 and the actual rain removal effect in Fig. 4 are obtained.

Table 1. Quantitative analysis of common data sets

Rain removal algorithms	Rain100L		Rain100H	
	PSNR	SSIM	PSNR	SSIM
Literature [18]	37.84	0.9814	37.84	0.9153
Literature [29]	31.21	0.9287	31.21	0.8654
Literature [30]	35.33	0.9694	35.33	0.9153

Table 2. Quantitative analysis of synthetic data sets

Rain removal algorithms	Rain2000L		Rain2000H	
	PSNR	SSIM	PSNR	SSIM
Literature [18]	37.83↓	0.9830↑	33.70↑	0.9449↑
Literature [29]	31.38↑	0.9627↑	27.75↓	0.8889↑
Literature [30]	35.31↓	0.9702↑	30.15↓	0.9342↑

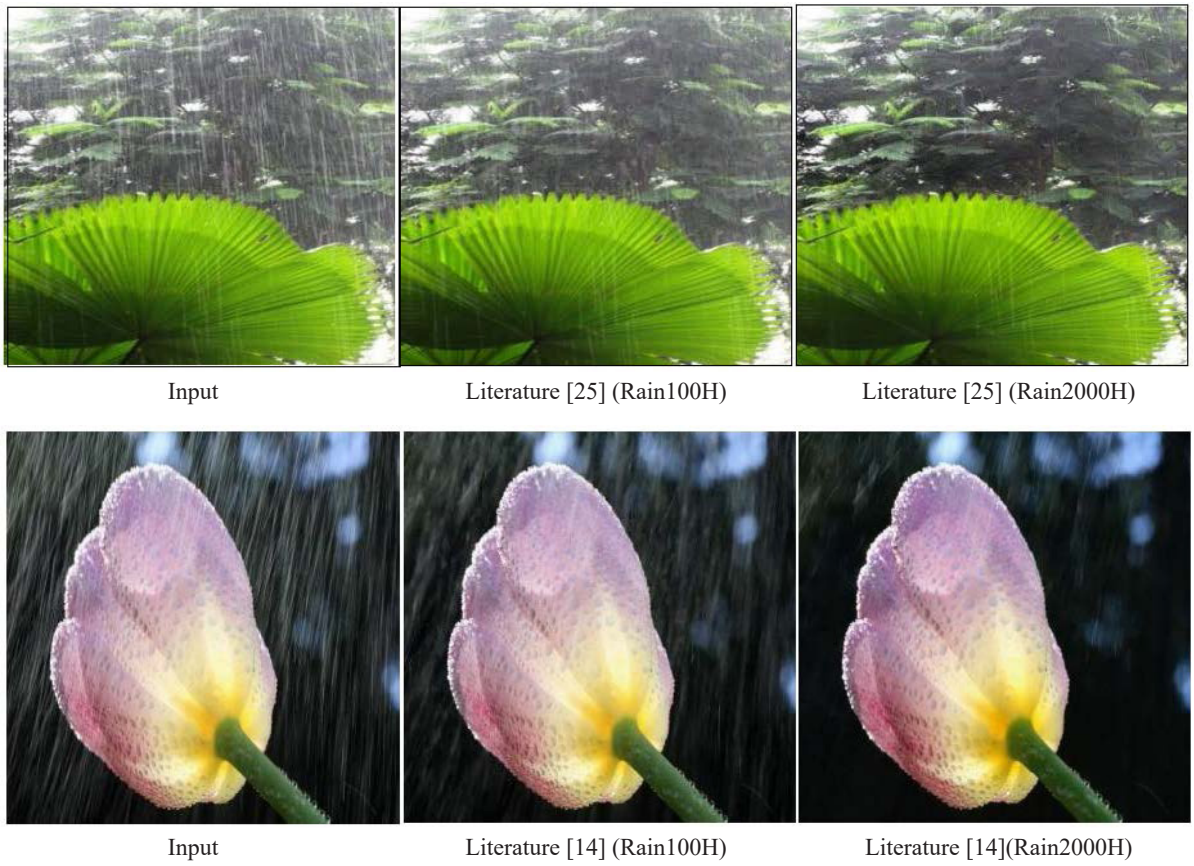


Fig. 4. The effect of removing rain in the actual pictures

In Table 1 and Table 2, the upward or downward trend is indicated by arrows. The experimental results suggest that the indicators generally have an upward trend, among which the PSNR has slightly decreased, with a maximum decrease of 0.07, while the SSIM has improved to some extent. At the same time, the actual rain removal effect in Fig. 4 is also significantly improved. It can be seen that the dataset constructed in this paper is to improve the structural similarity by reducing the peak signal-to-noise ratio, and the criti-cal information is the change of rain streak.

Fig. 4 shows the comparison of rain removal effects of various networks for the actual construction dataset in this paper. It can be seen from the actual results that in the process of removing rain from the real scene, various networks cannot effectively remove the rain streak information in most of the real rain for Rain_100H, but have remarkable results for Rain_2000H. It was analyzed that the real rainfall this time was mostly heavy rain with long rain streaks and the rain streaks are superimposed on each other. Consequently, the neural network does not learn the required rain features for rain removal in the real scenes well on Rain_100H dataset. By comparing the dataset proposed in this paper with the current popular datasets, it can be found that the rain streak information

is generally bright in the current popular datasets. These datasets only have the characteristics of rain and cannot reflect the real rain information. The synthetic rain method presented in this paper can be more suitable for the rain in the real environment, and also contains various rain streak information.

4.3 Optimization Network Comparative Experimental Results Analysis

Based on the above-constructed nightscape synthetic rain dataset Rain_dark100H, the performance of the MFR-Net proposed above is compared with multiple networks through the PSNR and SSIM indicators. The experimental results obtained are shown in Table 3.

In addition, Fig. 5 and Fig. 6 illustrate the comparison of the effects of different network models on real-world rain removal after training with synthetic datasets.

Table 3. The mean PSNR and SSIM of the composite dataset Rain_dark100H

	PSNR	SSIM
Literature [18]	30.02	0.8864
Literature [29]	30.47	0.8419
Literature [30]	27.23	0.8522
Literature [31]	26.79	0.8053
Ours	31.36	0.8978

It can be seen from the comparative experimental results that the algorithm designed in this paper can more effectively extract the scale features and shape features of the rain streak in the real rain removal scene, maximize the model performance, and reduce the transmission loss and gradient disappearance.

Fig. 5. Real rain image de-raining result

From Fig. 5 and Fig. 6, it can be seen that the obvious rain streaks have been removed, especially in Fig. 5 showing a clean background. The main drawbacks of Literature [18] in the comparison show that it fails to remove the rain streaks completely. Clearly, Literature [29] and Literature [30] have the ability to remove most

of rain streaks in different rain cases, but bring serious rain artifacts and blurred regions to the de-rained image. Some detailed information is lost in Literature [31] and leads color degradation to a certain extent. In general, the proposed MFR-Net can remove rain streaks and preserve color and detailed information, even in heavy rain conditions.

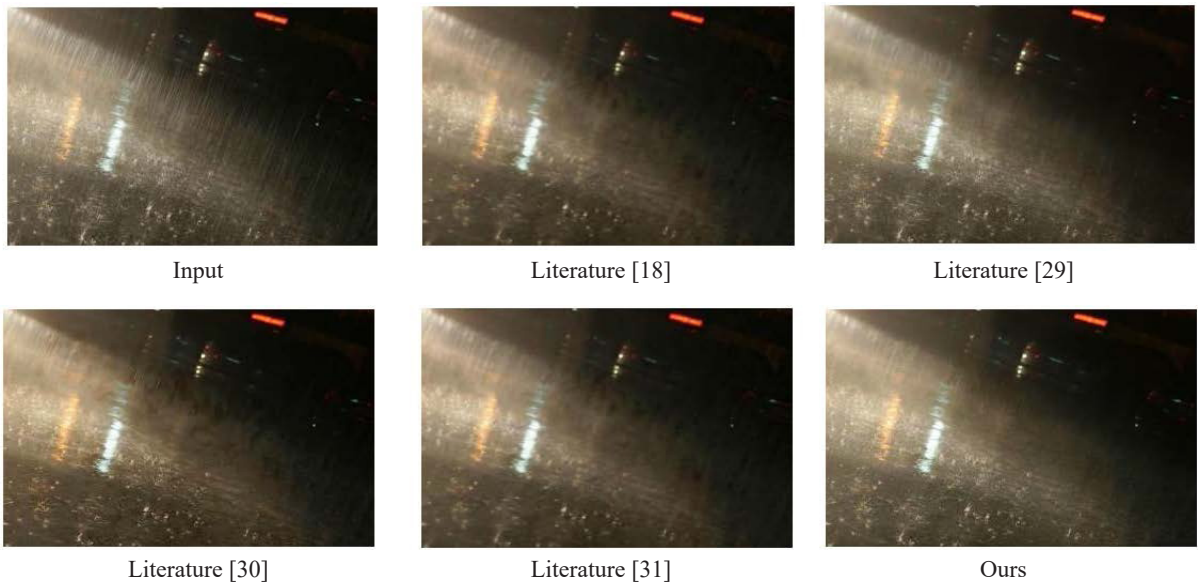


Fig. 6. Real rain image de-raining

However, blur and rain-like fog appear in the resulting image at the same time, which is one of the most common problems of image de-raining, especially in nightscapes. Through in-depth analysis of nightscape pictures, this paper makes a reasonable guess regarding the cause of the phenomenon: taking a photograph in nightscape requires a certain intensity of light source illumination, and only through the reflection of the light source can the camera capture the rain streak information. At the same time, for the raindrops that just stay on the light beam, after the rain removal is completed, the light beam information will still remain, resulting in the rain and rain-like fog phenomenon shown in the above picture.

4.4 Ablation Experiments

In order to verify the effectiveness of the improved network in removing rain from nightscape rain maps, an ablation experiment is carried out in this paper. All ablation experiments apply the Rain_dark100H dataset and are performed in the same operating environment. The main improvement of the MARD-Net in this paper is to increase its scale (the number of layers DS is set to 8 layers) and add the Global self-attention module. The baseline module is set as a Single-scale Attentive Residual Network and does not contain the GSA module, and then the GSA module, the deep-level scale DS and the superposition of the two are added respectively. The comparison results are depicted in Table 4.

Table 4. Ablation results

	PSNR	SSIM
Baseline	30.02	0.8864
GSA	30.67	0.8868
DS	30.43	0.8897
GSA+DS	31.36	0.8978

The experimental results show that for the same dataset, both PSNR and SSIM evaluations are improved by adding modules to the baseline module. Especially, the effect of the two modules working together is better than the effect of each module acting alone, which further proves the necessity and effectiveness of module fusion.

5 Conclusion

In this work, we make a study on the rain removal of nightscape rain maps. We exploit the characteristics of rain formation, and proposed a convenient method to build rainy images that complements the blank of nightscape rain image dataset. Then, we develop an end-to-end MFR-Net for rain images, which adopts a U-shape codec for exploring the features of rain streaks and follows by a GSA module to remove tiny rain streaks. Experimental results indicate that our proposed MFR-Net outperforms popular algorithms and effectively removes rain streaks while retaining the image details well. In future research, we will further strengthen the exploration of the features at different levels of the image. Meanwhile, we tried to find a solution to the problem of blurring and rain-like fogging of images after image de-rain.

References

[1] X. Wang, T. Xiao, Y. Jiang, S. Shao, J. Sun, C. Shen, Repulsion loss: Detecting pedestrians in a crowd, in: Proc. 2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018.
[2] J. Wen, H.-J. Wang, Y.-N. Li, J.-W. Zhao, Y.-Y. Yang, Multi-Update Patterns and Validity Verification for Robust Visual Tracking, Jour-nal of Computers 31(4)(2020) 77-90.
[3] Y.-C. Chen, Y.-Y. Lin, M.-H. Yang, J.-B. Huang, Show, match and segment: Joint weakly supervised learning of semantic matching and object co-segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 43(10)(2021) 3632-3647.
[4] W. Yang, R.-T. Tan, S. Wang, Y. Fang, J. Liu, Single image deraining: From model-based to data-driven and beyond, IEEE Transactions on Pattern Analysis and Machine Intelligence 43(11)(2021) 4059-4077.
[5] Y. Luo, Y. Xu, H. Ji, Removing rain from a single image via discriminative sparse coding, in: Proc. 2015 IEEE International Conference on Computer Vision, 2015.
[6] S. Gu, D. Meng, W. Zuo, L. Zhang, Joint convolutional analysis and synthesis sparse representation for single image layer separation, in: Proc. 2017 IEEE International Conference on Computer Vision, 2017.
[7] H. Zhang, V.-M. Patel, Convolutional sparse and low-rank coding-based rain streak removal, in: Proc. 2017 IEEE Winter Conference on Applications of Computer Vision, 2017.
[8] Y. Li, R.-T. Tan, X. Guo, J. Lu, M.-S. Brown, Rain streak removal using layer priors, in: Proc. 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016.
[9] M. Shao, L. Li, H. Wang, D. Meng, Selective generative adversarial network for raindrop removal from a single image, Neurocomputing 426(2021) 265-273.
[10] H. Lin, C. Jing, Y. Huang, X. Ding, A2Net: Adjacent aggregation networks for image raindrop removal, IEEE Access 8(2020) 60769-60779.
[11] L. Ji, C. Fu, W. Sun, Soft fault diagnosis of analog circuits based on a ResNet with circuit spectrum map, IEEE Transactions on Circuits and Systems I: Regular Papers 68(7)(2021) 2841-2849.
[12] R. Yasarla, V.-M. Patel, Uncertainty guided multi-scale residual learning-using a cycle spinning cnn for single image de-raining, in: Proc. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
[13] C. Wang, Y. Wu, Z. Su, J. Chen, Joint self-attention and scale-aggregation for self-calibrated deraining network, in: Proc. 2020 28th ACM International Conference on Multimedia, 2020.
[14] H. Wang, Q. Xie, Q. Zhao, D. Meng, A model-driven deep neural network for single image rain removal, in: Proc. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.
[15] Y. Wang, Y. Song, C. Ma, B. Zeng, Rethinking image deraining via rain streaks and vapors, in: Proc. 2020 European Conference on Computer Vision, 2020.
[16] C. Yu, Y. Chang, Y. Li, X. Zhao, L. Yan, Unsupervised image deraining: Optimization model driven deep cnn, in: Proc. 2021 29th ACM International Conference on Multimedia, 2021.
[17] Y. Nanba, H. Miyata, X. -H. Han, Dual Heterogeneous Complementary Networks for Single Image Deraining, in: Proc. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022.
[18] X. Chen, Y. Huang, L. Xu, Multi-scale Attentive Residual Dense Network for Single Image Rain Removal, in: Proc. 2020 Asian Confer-ence on Computer Vision, 2020.
[19] W. Yang, R. -T. Tan, J. Feng, J. Liu, Z. Guo, S. Yan, Deep joint rain detection and removal from a single image, in: Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[20] X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, J. Paisley, Removing rain from single images via a deep detail network,

in: Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[21] Y. Yan, W. Ren, Y. Guo, R. Wang, X. Cao, Image deblurring via extreme channels prior, in: Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017.
[22] Y. Du, J. Xu, Q. Qiu, X. Zhen, L. Zhang, Variational image deraining, in: Proc. 2020 IEEE/CVF Winter Conference on Applications of Computer Vision, 2020.
[23] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.-N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, in: Proc. Advances in Neural Information Processing Systems 30, 2017.
[24] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, An image is worth 16x16 words: Transformers for image recognition at scale, <https://arxiv.org/abs/2010.11929>, 2020 (accessed 03.06.20).
[25] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: Proc. 2021 IEEE/CVF International Conference on Computer Vision, 2021.
[26] Z. Shen, I. Bello, R. Vemulapalli, X. Jia, C.-H. Chen, Global self-attention networks for image recognition, <https://arxiv.org/abs/2010.03019 >, 2020 (accessed 14.10.20).
[27] Y. Mei, Y. Fan, Y. Zhang, J. Yu, Y. Zhou, D. Liu, Y. Fu, T.S. Huang, H. Shi, Pyramid attention networks for image restoration, <https://arxiv.org/abs/2004.13824 >, 2020 (accessed 03.06.20).
[28] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Proc. 2015 International Conference on Medical Image Computing and Computer-Assisted Intervention, 2015.
[29] N. Jiang, W. Chen, L. Lin, T. Zhao, Single image rain removal via multi-module deep grid network, Computer Vision and Image Under-standing 202(2021) 103106.
[30] T. Wang, X. Yang, K. Xu, S. Chen, Q. Zhang, R.-W. Lau, Spatial attentive single-image deraining with a high quality real rain dataset, in: Proc. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019.
[31] K. Jiang, Z. Wang, P. Yi, C. Chen, B. Huang, Y. Luo, J. Ma, J. Jiang, Multi-scale progressive fusion network for single image deraining, in: Proc. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020.