Multi-Stage Enhanced Rain Image Restoration Network

Quanqing Wang1, Yuanshun Cheng2, Yueting Yang3, Chao Wang1,4 and Chuansheng Yang1,4*

1School of Information Engineering, Zhejiang Ocean University, Zhoushan, Zhejiang, 316022, China
2Tianfu College of SWUFE, Mianyang, Sichuan, 621000, China
3School of Mathematics and Statistics, Beihua University, Jilin, Jilin, 132013, P.R. China
4Key Laboratory of Oceanographic Big Data Mining and Application of Zhejiang Province, Zhoushan, Zhejiang, 316022, China
*Corresponding author’s e-mail: cshyang@163.com

Abstract. The restoration task of rain images is beneficial to the further development of computer vision tasks. However, recent rain removal methods have difficulty coping with dense rain streaks. In this paper, we propose a Multi-stage Enhanced Rain Image Restoration Network (MERN) for the task of rain removal. Our main idea is to design a three-stage network architecture, which makes full use of the enhanced information of the previous stage and complements the information of the next stage through transmission step by step. We give consideration to information enhancement at the same time of information transmission, fully improve the information utilization rate and the overall network effect.

1. Introduction
In rainy weather, the cover of background information by raindrops and rain streaks is a factor that cannot be ignored, which will greatly reduce the performance of further computer vision tasks such as recognition and classification. Rain removal for a single image is to recover a clean background image from an image affected by rain streaks degradation. Removing dense and complex rain streaks is still a task full of great challenges, while a series of relevant methods have been proposed to effectively remove rain streaks in light rain scenes.

1.1. Related Works
A large number of rain-removal methods rely on the following formula:

\[ I = J + \sum_{i}^{n} S_i \]  

where \( I \) is the observed input image, \( J \) is the background scene free from the rain streaks. And \( S_i \) represents the rain layer, with \( n \) as the total number of rain streaks layers. In general, methods for image derain can be divided into two categories: One is based on the physical models and the other is data driven.

Early rain removal adopts physical model-based approaches, utilizing various prior constraints, e.g., low-rank [1], sparse code [2], and Gaussian mixture model [3]. Luo et al. [2] decomposed rain streaks and clean backgrounds based on a new discriminative sparse coding framework. Li et al. [3] applied the
Gaussian mixture model on the patch of rain streaks and clean backgrounds to remove rain streaks. However, the general effect of the physical based rain removal method is not good enough and it cannot meet the requirements of computer vision.

In recent years, the rapid development of deep neural networks brings prospects for rain removal. Correlational studies [4,5,6] have adopted the method of supervised learning and achieved a very significant rain-free effect, while there are also drawbacks. On the one hand, many current methods usually only estimate rain streaks images or rain-free images [4,5], ignoring the enhanced effect brought by the combination of the two ways. On the other hand, many deep learning-based methods [4,6] do not consider the difference of noise in the extraction of rain streaks, leading to the imperfection of rain streaks removal. Rain removal methods have achieved good results, but there are still major limitations, such as the inability to obtain clear background images for dense rain streaks.

1.2. Motivation

In this paper, we propose a multi-stage enhanced rain image restoration network (MERNet) for the task of rain removal. Since a multi-stage structure can divide the restoration into more manageable steps, we introduce a progressive approach to gradually learn the recovery features of degraded inputs. We innovatively employ addition and multiplication networks [7] to multi-stage tasks and refine extraction of different noises to make up for the information that U-Net could not obtain. Additive and multiplicative noise estimation can effectively improve the fine degree of the extracted rain streaks, so as to restore the better residual components.

Our contributions can be summarized as follows:

- We propose a new Multi-Stage Enhanced Rain Image Restoration Network (MERNet) based on [8]. The network can remove the dense rain streaks on the images and enhance the restoration effect of background details.
- Based on the analysis of the additive and multiplicative noise, we innovatively use additive and multiplicative network in the multi-stage rain removal task to better enhance the rain removal effect.
- In order to better guide the recovery effect, we used the deep feature-guided additive and multiplicative network to analyze and process the rain images.

2. Proposed Method

In this section, we will specifically introduce MERNet whose architecture can effectively remove rain streaks and restore background details.

2.1. Overall network structure

As shown in figure 1, we propose an end-to-end multi-stage rain removal network, which mainly consists of three processing phases. For the first and second stages, we use the U-Net structure to extract the context features and innovatively employ additive and multiplicative networks to generate predicted residual components as they can focus on different perspectives of noise. Then a fusion module is used to fuse the rough rain-free images predicted by the three sub-networks to generate a clean image of the corresponding stage. In the last stage, the channel attention module is used to process the original image and enrich the spatial details to the greatest extent. MERNet improves the utilization rate of information, carries out information enhancement from a multi-perspective, and improves the recovery effect of rain images.
2.2. Network details

Based on the structure of [8], we integrate additive and multiplicative networks guided by deep feature information, which is extracted from the second layer of the encoder. The predicted residual factors are added or multiplied by the original input image to realize the reconstruction of the clean image. Through the residual channel attention fusion module, we selectively fuse the three rough rain-free images, which are respectively processed by U-Net, additive sub-network, and multiplicative sub-network, to get the rain-free images of Stage 1 and Stage 2.

The concrete structure of additive sub-network (AS) and multiplicative sub-network (MS) are shown in figure 2. AS is composed of two multi-scale convolution modules, i.e., Residual Inception Modules (RIM), and is used to predict additive residual factors. The combination of multi-scale convolution blocks can expand the receptive field, enhancing the ability of the network to capture rain streaks. With the help of RIMs, AS analyzes the encoding feature and learns the rain streaks of different shapes and directions in the rain images. The AS also has better help for the restoration of high-frequency details and the dark part of the rain streaks removal.

MS focuses on channels with richer information and estimates multiplicative residual factors in the channel. MS contains two Channel-attentive Inception Modules (CIM) which are also composed of multi-scale convolution block and channel attention block. The sub-network mainly focuses on the bright area and realizes the reconstruction of clean images from the perspective of multiplicative noise.
Figure 2. Network structure of addition network AS and multiplication network MS. (a) is the structure of AS. (b) is the structure of MS.

Instead of predicting the restored image directly, we add and multiply the predicted residual factors to the input rain map to get clean images. In order to improve the rain-removal effect, we incorporate additive and multiplicative networks into the three-stage progressive network. The MERNet uses the U-Net, addition and, multiplication networks to make preliminary predictions on images, and combine the attention mechanism to enhance useful information.

2.3. Loss function

We divided the input image into four blocks in the first stage, two blocks in the second stage, and use the original image in the last stage. We optimize the end-to-end MERNet with the following loss function:

\[ L = \sum_{s=1}^{3} [L_{\text{char}}(X_s, Y) + \lambda L_{\text{edge}}(X_s, Y)], \]  

where \(Y\) represents the ground-truth image, and \(L_{\text{char}}\) is the Charbonnier loss [9] as follows:

\[ L_{\text{char}} = \sqrt{||X_s - Y||^2 + \epsilon^2} \]  

with constant \(\epsilon\) empirically set to \(1 \times 10^{-3}\) for all the experiments. In addition, \(L_{\text{edge}}\) is the edge loss defined as:

\[ L_{\text{edge}} = \sqrt{||\Delta(X_s) - \Delta(Y)||^2 + \epsilon^2}, \]  

where \(\Delta\) denotes the Laplacian operator. The parameter \(\lambda\) as in equation (2) controls the relative importance of the two loss terms, which is set to 0.05 as in [5].

3. Experimental Results

In this section, we will cover the experimental setup in detail. We compare the three most advanced rain removal methods. These methods are tested on Rain100H [10], Rain100L [10], and Test2800 [4]. PSNR and SSIM indicators are used for quantitative evaluation. The models are trained on 256 \times 256 patches with a batch size of 1 for \(1 \times 10^5\) iterations. We use Adam optimizer with the initial learning rate of \(2 \times 10^{-4}\) using the cosine annealing strategy [11]. All the experiments use PyTorch 1.7.0 on an Nvidia GeForce RTX3090 (24G). The quantitative experimental results are shown in table 1.

| Methods       | Rain100H | Rain100L | Test2800 |
|---------------|----------|----------|----------|
|               | PSNR     | SSIM     | PSNR     | SSIM     | PSNR     | SSIM     |
| RESCAN [6]    | 26.36    | 0.786    | 29.80    | 0.881    | 31.29    | 0.904    |
| MSPFN [5]     | 28.66    | 0.860    | 32.40    | 0.933    | 32.82    | 0.930    |
| MPRNet [8]    | 30.41    | 0.890    | 36.40    | 0.965    | 33.64    | 0.938    |
| MERNet (Ours) | 31.13    | 0.905    | 36.63    | 0.973    | 33.71    | 0.939    |
Compared with PSNR and SSIM of the other 3 methods, the value is proportional to the qualitative. Red is the highest and blue is the second highest.

Qualitative analysis is shown in figure 3. The results obtained by these methods are qualitatively compared on some rainy image samples. In terms of visual effects, our results are closer to the ground truth compared with the other five methods.

Through a series of experiments, it is shown that our network MERNet has the best effect of rain removal visually. The values of PSNR and SSIM have also been effectively improved.

4. Conclusion
MERNet, a multi-stage attention enhancement rain removal network, adopts flexible design principles and makes full use of the guidance effect of additive and multiplicative sub-network. The network can not only remove rain streaks but also restore the background details well. Extensive experiments show that the newly proposed model performs well on the rain removal task.

Acknowledgments
This work is supported by the Science and Technology Development Project Program of Jilin Province (20190303132SF); The Key Project of Natural Science Foundation Joint Fund of Jilin Province (2020122367JC); The National Natural Science Foundation of China (12001005).

References
[1] Chen, Y. L., & Hsu, C.T. (2013) A generalized low-rank appearance model for spatio-temporally correlated rain streaks. In Proceedings of the IEEE International Conference on Computer Vision. pp. 1968-1975.
[2] Luo, Y., Xu, Y., & Ji, H. (2015) Removing rain from a single image via discriminative sparse coding. In Proceedings of the IEEE International Conference on Computer Vision. pp. 3397-3405.
[3] Li, Y., Tan, R. T., Guo, X., Lu, J., & Brown, M.S. (2016) Rain streak removal using layer priors. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2736-2744.
[4] Fu, X., Huang, J., Zeng, D., Huang, Y., Ding, X., & Paisley, J. (2017) Removing rain from single images via a deep detail network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3855-3863.
[5] Jiang, K., Wang, Z., Yi, P., Chen, C., Huang, B., Luo, Y., ... & Jiang, J. (2020) Multi-scale progressive fusion network for single image deraining. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8346-8355.
[6] Li, X., Wu, J., Lin, Z., Liu, H., & Zha, H. (2018) Recurrent squeeze-and-excitation context aggregation net for single image deraining. In Proceedings of the European Conference on
[7] Vu, D.T., Gonzalez, J.L., & Kim, M. (2021) Exploiting Global and Local Attentions for Heavy Rain Removal on Single Images. arXiv preprint arXiv:2104.08126. https://arxiv.org/abs/2104.08126.

[8] Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.H., & Shao, L. (2021) Multi-stage progressive image restoration. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14821-14831.

[9] Charbonnier, P., Blanc-Feraud, L., Aubert, G., & Barlaud, M. (1994, November) Two deterministic half-quadratic regularization algorithms for computed imaging. In Proceedings of 1st International Conference on Image Processing., Vol.2. pp. 168-172.

[10] Yang, W., Tan, R.T., Feng, J., Liu, J., Guo, Z., & Yan, S. (2017) Deep joint rain detection and removal from a single image. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1357-1366.

[11] Loshchilov, I., & Hutter, F. (2016) Sgdr: Stochastic gradient descent with warm restarts. arXiv preprint arXiv:1608.03983. https://arxiv.org/abs/1608.03983.