

PAPER • OPEN ACCESS

Rain Removal from images based on Computer Vision and conditional Generative Adversarial Networks

To cite this article: Yunfei Li *et al* 2020 *J. Phys.: Conf. Ser.* **1570** 012064

View the [article online](#) for updates and enhancements.

You may also like

- [Fast simulation of muons produced at the SHiP experiment using Generative Adversarial Networks](#)
C. Ahdida, R. Albanese, A. Alexandrov et al.
- [Face Aging on Realistic Photo in Cross-Dataset Implementation](#)
Nabilah Hamzah and Fadhlan Hafizhelmi Kamaru Zaman
- [Ultra-fast generation of air shower images for Imaging Air Cherenkov Telescopes using Generative Adversarial Networks](#)
Christian Elflein, Stefan Funk and Jonas Glombitza



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Rain Removal from images based on Computer Vision and conditional Generative Adversarial Networks

Yunfei Li¹, Lidan Wang^{1,3,4,5}, Taixing Chen¹, Ziyuan Wang¹, and Shukai Duan^{1,2,3,4,5,*}

¹College of Electronic and Information Engineering, Southwest University, Chongqing, 400715, China

²College of Artificial Intelligence, Southwest University, Chongqing, 400715, China

³Brain-inspired Computing & Intelligent Control of Chongqing Key Lab, Chongqing, 400715, China

⁴National & Local Joint Engineering Laboratory of Intelligent Transmission and Control Technology, Chongqing, 400715, China

⁵ChongqingBrain Science Collaborative Innovation Center Chongqing, 400715, China

*Corresponding author's e-mail: duansk@swu.edu.cn

Abstract. In view of the quick development of modern computer technology, more pictures of natural scenes have been collected and analyzed for application in military fields, scientific research, and computer vision fields related to object detection. However, the sharpness of these pictures collected will be affected by rain and snow, which will reduce the visibility of things in the pictures. This will not only affect the viewing, but also affect scientific research fields such as autonomous driving and video surveillance. Therefore, the research on image rain removing is very practical. An improved algorithm model named pixel-GAN is proposed in this article, which based on conditional Generative Adversarial Networks (cGAN). This new method regards the problem of removing rain from specific images as the style conversion problem. What we improved is that a VGG-16 module is added to the original network to extract perceptual loss as a new part of the loss function. We show the excellent rain-removing ability of the improved algorithm and also illustrate the reason why we add the perceptual loss. Compared with other existing methods, there is no need to set a prior model, the training speed is faster and the proposed model finally generates higher quality pictures.

1. Introduction

Now that we learn about the significance of removing rain from specific images in the field of various production researches including object detection and some high-precision scientific researches. Therefore, it is extremely important to find effective ways to remove rain from pictures. At present, there are several ideal methods for removing rain: 1. Rain removal based on sparse coding and classifier. 2. A method of mathematical derivation based on a physical model. 3. Removing rain from single images via a deep detail network [1]. 4. Deep joint rain detection and removal from a single image [2]. The first two methods need to create prior model and the last two methods implemented are on the basis of convolutional neural network.



The methods already proposed perform well in the work of removing rain. Compared with the first two methods, the one using convolutional neural network performs better. It trains models faster, and overcomes some of the limitations of manually designing features and models. Although the results are excellent, there are still some areas for improvement in these studies. For example, the lack of details in the pictures needs to be caught and the color of images can be slightly distorted. This article implements an end-to-end training method by modifying the model proposed in pix2pix [3] for rain removing. After adding the perceptual loss, the model can better ensure the consistency of the content between the corresponding images and the images after rain removal get closer to real clear images.

Paired pictures are trained in this work, and the VGG-16 module is added to the original generated network to measure the difference in the last feature layers before and after image processing. The model finally showed superior ability to remove the rain in the pictures. Details will be introduced in the next chapters. In the second chapter we mainly introduce the related work of the article and the proposed model and related loss function will be introduced in the third chapter. The fourth chapter includes the main experiments and we make a summary in the last chapter.

2. Related work

With the rapid development of deep convolutional neural network in the field of computer vision, more and more researchers apply it to the field of image restoration. Several methods with better results have been listed. The first two methods based on pure physical models involve a lot of physics expertise, so no specific introduction be made here. Let's focus on methods using convolutional neural networks. The deep convolutional neural network combining rainwater detection and removal aims to achieve the separation of rain layer and background layer. The deep architecture includes a novel network structure, a context-expanded network based on contextual information. This structure is used to extract the identifiable features of the rainwater image and to serve as a basis for subsequent detection and removal. The idea for this article comes from the method for single image haze removal [4] and image-style transfer using convolutional neural networks [5] these two articles proposed. Pixel-GAN is used to solve this problem. We regard the rain-removal problem as a problem of image style conversion and simplified network structure compared to the above model. The generator learns the mapping relationship between the paired pictures, reducing the setting of the rain layer model.

The earliest Generative Adversarial Networks [6] (GAN) consists of a discriminator and a generator, which proposed by Ian.J. Goodfellow on the article named Generative Adversarial Networks in 2014. A complete generative adversarial network structure is shown in Figure 1.

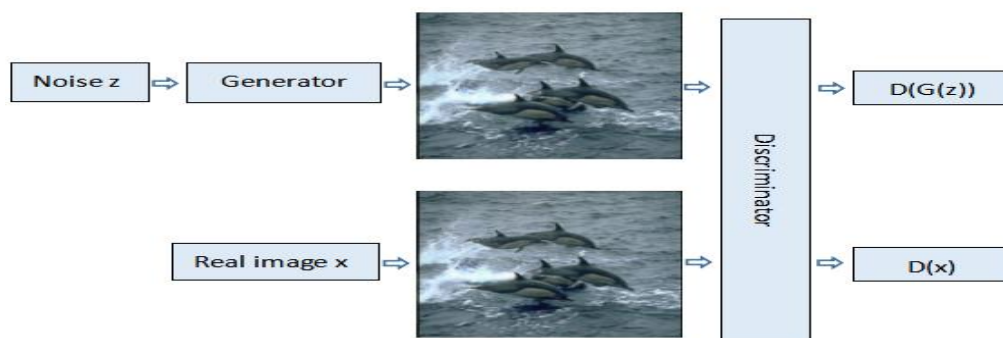


Figure 1. The structure of Generative Adversarial Network

A noise z as the input of the generator as shown in the Figure 1, and the generator learn the feature distribution of the real picture to generate realistic samples. Then the discriminator judges whether the input sample comes from the generated data or not, and outputs the corresponding probability value, which represents the probability that the picture comes from the real sample. In the course of confrontation, the two parts constantly optimize their own networks and parameters. Finally, the generator can't determine the true source of the input picture. At this time, the discriminator is abandoned and the generator can be used to generate the picture we need.

However, the style of the image generated by such a generator is non-directional, and the cGAN proposed later solves this problem. By adding labels to the generator and discriminator, various of sample images can be generated as designers desired. Pix2pix is a special cGAN model, in which the original noise is removed and the network is directly trained with paired pictures. The model has achieved remarkable results in the field of image processing and has gradually applied to other image translation work.

3. Network structure

In the experiment, a VGG-16 module is added to the original model structure to extract the perceptual loss. The structure of VGG-16 module is shown in Figure 2:

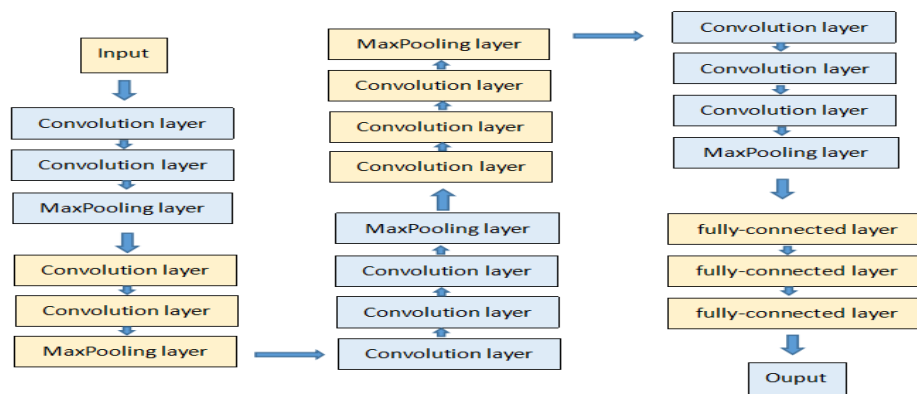


Figure 2. The network structure of VGG-16

The VGG-16 module includes eight parts. The extracted perceptual loss is added to the loss function as a new term, which further ensures the consistency of content between the generated image and the original image. The new structure of model with VGG-16 module can be seen in Figure 3

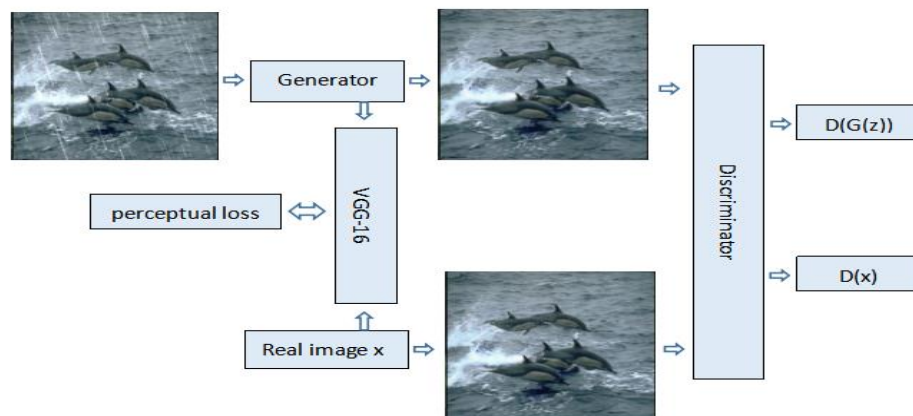


Figure 3. The network structure of the new model proposed

The U-net network is used in this model, which was proposed in 2015 and initially applied to the analysis of some medical images. It is a symmetric convolutional neural network structure, including three parts: encoder, converter and decoder, forming a symmetrical structure. In order to retain some information shared between the input and output layers, jumper connections are added to the last few layers of the network, that is, the symmetrical two-layer characteristic channels connected together.

The discriminator uses a convolutional "Patch-GAN" classifier. This classifier penalizes the structure at the scale of the patch. The Patch-GAN model here divides the generated picture into blocks of size $N \times N$, and inputs the probability value of the judgement given by the discriminator. We

take the average of the probabilities of all local pictures as the final score. The processing of the local information of the picture in this way helps to restore the loss of details in the parts blocked by rain.

The original loss function of pix2pix is

$$G = \arg \min_G \max_D L_{cGAN}(G, D) + \lambda L_{L1}(G) \quad (1)$$

Among them

$$L_{cGAN}(G, D) = E_{x, y \sim P_{data}(x, y)} [\log D(x, y)] + E_{x \sim P_{data}(x), z \sim P_z(z)} [\log 1 - D(x, G(x, z))] \quad (2)$$

Here the goal of generator is to minimize this function value, and the discriminator aims to maximize this function value. In the figure, we added the VGG-16 module to add the Lp loss, so the final loss function is

$$G^* = L_{GAN} + \lambda_1 L_1 + \lambda_2 L_p \quad (3)$$

Formula three is the improved loss function we proposed. This method takes the average of perceptual loss of the last three feature layers as the final perceptual loss.

4. Experiment

The complete experiments performed on the rain12 dataset. In the experiment, we take the average of the perceptual loss of the last three feature layers as the final perceptual loss. Two thousand sets of rain and clear matching pictures are trained in the training set and the test set use a hundred sets of pictures. Alternate training of generators and discriminators is used in the experiment. We set the batchsize in training to 16. After training two hundred epochs on the data in the training set, we evaluate the quality of pictures generated from both qualitative and quantitative assessments.

Qualitative evaluation: Figure 4 and Figure 5 shows the experimental results after different training epochs, and the comparison of the experimental results before and after adding the perceptual loss is shown in Figure 6.



Figure 4. The experimental results of the dataset after training 50 epochs



Figure 5. The experimental results of the dataset after training 200 epochs.

The experimental results indicate that the ideal experimental results have been generated after training 50 epochs of the dataset, and the generate effect increases slightly after training for 200 epochs, which shows that the loss function after adding the perceptual loss can also be quickly weakened.

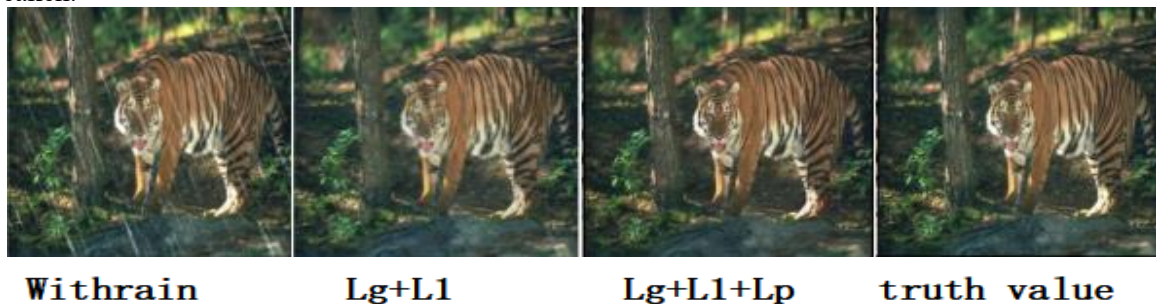


Figure 6. Different losses induce different quality of results. Here are the results under different losses.

The experimental results before adding perceptual loss also show better rain removal ability, but there will be a slight color deviation. This problem has been successfully solved by using the new loss function.

Quantitative evaluation: The structural similarity (SSIM) and peak signal-to-noise ratio (PSNR) two main measures are mainly used to evaluate the effect of rain removal by different methods. The larger the corresponding metric value is, the better the effect performs.

Baseline	Rain12	
Metric	PSNR	SSIM
DSC	30.02	0.87
LP	32.02	0.91
CNN	26.65	0.78
SRCNN	34.41	0.94
Pixel-GAN	35.21	0.96

Figure 7. Results of different rain removal methods in PSNR and SSIM metrics

In terms of PSNR and SSIM, the metric values of pictures generated by the network proposed in this article have achieved good results. Compared with the previous better methods, the new model exhibits superior image recovery ability. This also reflects the protruding capabilities of Pixel-GAN in image generation.

5. Conclusion

A new end-to-end network structure model is used to remove rain lines from pictures in this article. In the experiment, the symmetrical U-net structure of the generator is used to retain the underlying features of the picture before and after the rain removing. The Patch-GAN in the discriminator is used to learn the details, for better processing the local information. Besides, the perceptual loss added can ensure the restoration of the picture content consistent on the image, which helps to achieved remarkable results in the work of removing rain from pictures. However, the performance of removing raindrops on the lens is not very good, because the raindrops are too large to recovered, and the location of the raindrops is random. In the latest method to remove raindrops, an attentive network [7] is added to detect raindrop positions, which increases the processing of local information. In subsequent studies, an attention network is considered to combine with our experiment. The combination of attention network and Patch-GAN may achieve better results in the recovery of local information

Acknowledgments

The work was supported by the National Key R&D Program of China (Grant nos. 2018YFB1306600), the National Natural Science Foundation of China (Grant nos. 61571372, 61672436, 61601376), the Fundamental Research Funds for the Central Universities (Grant nos. XDJK2016A001, XDJK2017A005), the China Postdoctoral Science Foundation Special Funded (2018T110937), the Chongqing Postdoctoral Science Foundation Special Funded (Xm2017039), and the Fundamental Science and Advanced Technology Research Foundation of Chongqing(cstc2017jcyjBX0050, cstc2016jcyjA0547).

References

- [1] Xueyang Fu, Jiabin Huang, Delu Zeng. (2017) Removing rain from single images via a deep detail network. In: The IEEE Conference on Computer Vision and Pattern Recognition. Hawaii.
- [2] Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, Shuicheng Yan. (2017) Deep joint rain detection and removal from a single image. In: The IEEE Conference on Computer Vision and Pattern Recognition. Hawaii.
- [3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. (2017) Image-to-Image Translation with Conditional Adversarial Networks. In: The IEEE Conference on Computer Vision and Pattern Recognition. Hawaii.
- [4] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, Dacheng Tao. (2016) Dehazenet: An end-to-end system for single image haze removal. IEEE Transactions on Image Processing 25.11 : 5187-5198.
- [5] L.A. Gatys, A.S. Ecker, and M. Bethge. (2016) Image style transfer using convolutional neural networks. In: The IEEE Conference on Computer Vision and Pattern Recognition. Las Vega.
- [6] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014) Generative adversarial nets. In NIPS.
- [7] Qian R, Tan R T, Yang W, et al. (2018) Attentive Generative Adversarial Network for Raindrop Removal from a Single Image. In: The IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City.