Image Raindrop Removal Method for Generative Adversarial Network Based on Difference Learning

Renhe Chen¹, Zhenyi Lai¹ and Yurong Qian*¹

¹ School of Software, Xinjiang University, Urumqi, 830046, China
*Corresponding author’s e-mail: qyr@xju.edu.cn

Abstract. Due to the interference of the external environment such as rainy weather on the camera, raindrops can easily adhere to the lens and seriously affect the quality of the photos taken. Therefore, it is of great significance to remove raindrops from the image and improve the quality of the photo. In this paper, a raindrop method for generative adversarial network images based on differential learning is proposed. The general generative network is to input images with raindrops and output clean images. The generative network in this paper does not directly output clean images, but learning the difference between images with raindrops and without raindrops, then subtract the learned difference from the image with raindrops to generate a clean image. In order to learn this difference more effectively, adding reconstruction loss to the generative network, the pre-trained VGG-16 network is used to extract the difference between the generated image and the real image features and calculate the mean square error. The experimental results show, the method in this paper can not only remove the raindrops in the image well, but also reconstruct the image information of the part blocked by the raindrops. The image processed by the algorithm in this paper is tested using the yolov3 target detection algorithm, which can significantly improve the recognition accuracy of the detection algorithm.

1. Introduction

When taking photos of rainy days and other scenes, it is easy to be infected with some raindrops on the lens, which will also be taken into our photos, resulting in the decline in the quality of the image, which not only leads to the deterioration of the subjective visual effect, but also seriously affects the recognition accuracy of the object detection and tracking algorithm of computer vision[1]. Therefore, the effective removal of these raindrops can not only improve the subjective visual effect, but also provide high-quality images in the subsequent processing of the computer vision system and improve the performance of the computer vision system.

At present, raindrop removal methods can be divided into two categories: one is based on traditional models, and the other is based on deep learning. Traditional methods use PCA to learn the shape of raindrops and match one region of the test image to the region of the raindrops being studied [2]. However, because of the transparency and shape anisotropy of the raindrop, PCA cannot well simulate the various appearance of raindrop. Yamashita et al. [3] used a stereo system to detect and remove raindrops by comparing the difference in stereo measurements with the distance from the stereo camera to the lens surface. Although this method can remove raindrops to some extent, it cannot be directly applied to an image because there is no information about the distance between camera and lens in the ordinary single image.
In recent years, many scholars have adopted deep learning to remove raindrops. Because of the blocking of raindrops, the background information of the image will be blurred and lost. The key goal is to remove the raindrops and reconstruct the missing image. Eigen et al. [4] used convolutional neural network to remove raindrops, which used raindrop images and clean images as training sets for learning and training. This method has a good effect on images with less raindrop, but for image processing results with dense raindrop, raindrop residues and image blurring phenomena will occur. Remove raindrops and rebuild the lost image information can be defined as the image translation tasks, generative adversarial network[5] in many images translation work has obtained the good effect, Pix2Pix to convert an image into another image, it uses the generated adversarial learning input image of rain to the output clean image, this is a commonly used method[6]. This method can well solve the problem of image blur, but there will still be raindrop residue when dealing with large area raindrops.

This paper proposes a generative adversarial network based on difference learning, the generative network does not directly learn images with raindrop to clean images, because the image can be any complex scenes, as a result, the generative network needs to learn more scenes to reconstruct clear images, but the distribution of raindrop relatively complex scene is relatively single. The generative network can better learn the distribution of raindrops. Therefore, the generative network of this paper learns the difference between images with raindrop and clean raindrop, and then the raindrop image is obtained by subtracting the difference from the raindrop image.

2. Proposed raindrop removal method

Figure 1 shows the general framework of the proposed network model. The basic model of generative adversarial network is adopted. The model of raindrop removal can be regarded as I=C+R, and the image with raindrop (I) is composed of the clean image (C) and raindrop (R), and raindrop (R) is the difference between the image with raindrop and the clean image. There are two main networks in the model, the generative network and the discriminative network. The generative network input image with raindrop (I), trained to output raindrop image(R). The discriminative network will evaluate whether the difference images generated are real.

![The generative network](image1)

![the discriminative network](image2)

Figure 1. The general framework of the model

The general loss function in this paper can be expressed as:

$$\min G \max D V(D, G) = E_{x \sim P_{data}} [\log(D(x))] + E_{x \sim P_{G}} [\log(1 - D(G(x)))]$$

(1)

Where G represents generative network, D represents discriminative network, $p_{data}$ represents image label without raindrop, and $P_G$ represents image data set with raindrop.

2.1. The generative network

This paper adopts a generative network based on context auto-encoder. The purpose of generative network is to generate the distribution of raindrop image. The input of network is the image with
raindrop, which is composed of 16 convolutional layers, active layers, and skip connections are added to prevent the loss of image information. The convolution kernel is $5 \times 5$, and the last layer is the tanh activation function. As shown in Figure 2.

The loss function of the generative network adopts the reconstruction loss, which measures the overall difference between the raindrop image feature and the label feature of the real raindrop image produced by the generative network. The features are extracted by the convolutional neural network. In this paper, VGG-16 convolutional neural network structure is used to extract features [7]. The reconstruction loss function can be expressed as:

$$L_G = L_{MSE}(VGG(g), VGG(r))$$

(2)

Where VGG represents the convolutional neural network that has been pretrained in advance, $g$ represents the raindrop image distribution output by the generative network, and $r$ represents the image distribution of real raindrop.

### 2.2. The discriminative network

In order to distinguish the distribution of the generated raindrop image from the real raindrop image, the discriminative network uses the full convolution network to extract the image features. A typical discriminator only needs to output one value, true or false, to evaluate the entire image. The discriminator in this paper outputs an $N \times N$ matrix, in which there is only one true or false value for the element, and each value represents a relatively large perceptive field in the original graph, each element is a region block in the original graph. Compared with the method of evaluating the whole image with only one value, the output of $N \times N$ matrix can better train the details of the image and prevent the generation of fuzzy image. As shown in figure 3, the discriminative network is composed of 6 convolutional layers plus the relu activation layer, and the last layer is the sigmoid activation function. The convolution kernel is $5 \times 5$.

The loss function of the discriminator can be expressed as:

$$L_D = -\log(D(g)) - \log(1 - D(r))$$

(3)
Where \( g \) represents the distribution of the generated raindrop image and \( r \) represents the distribution of the real drop image.

3. Experimental Result

The experimental data set in this paper adopts the data set provided by Peking University. There are 1119 pairs of images with and without raindrop in the data set. We take one of the 800 pairs as a training set, the remaining as a test set, because this is the generative of network learning with the difference between image with raindrop and clean image, and does not give the data set with images with raindrop and clean image difference, in order to get the image of the differences label, the difference label is obtained by subtracting the image with raindrops and clean image, and resize to \( 240 \times 360 \) as the training sample.

The hardware configuration of the experiment is Intel Core i7, NVIDIA GeForce1060, 16GB of computer memory, window10 64bit, CUDA10.0, python3.6, and Google's deep learning framework TensorFlow1.12.0.

The contrast methods used in this paper is Eigen [4] and Pix2Pix [6], and the evaluation indexes include quantitative image evaluation method peak signal to noise ratio (PSNR), structural similarity (SSIM) and subjective judgment.

| methods        | PSNR  | SSIM  |
|----------------|-------|-------|
| Eigen          | 28.59 | 0.6726|
| Pix2Pix        | 30.14 | 0.8299|
| The proposed method | **31.66** | **0.8921** |

As can be seen from table 1 that the method proposed in this paper is superior to other algorithms in RSNR and SSIM, which indicates that the method proposed in this paper can generate more real results.

![Figure 4](image4.jpg)

Figure 4. (a) Image with raindrops. (b) Clean image. (c) Result image with raindrops of the proposed method. (d) Result clean image of the proposed method

In figure 4, from the brighter part of figure (c), it can be seen that the generative network finds out the distribution image of raindrops very well. The generated clean image keeps the details of the image while removing the raindrops, and recovers the information of the part blocked by raindrops. Improved visual perception compared to the original image with raindrops.

![Figure 5](image5.jpg)

Figure 5. (a) Image with raindrops. (b) Eigen. (c) Pix2Pix. (d) The proposed method

As can be seen from figure 5, Eigen method does not remove large area raindrops well. Although Pix2Pix removes raindrops, it still leaves shadows on the image, affecting the visual perception. The proposed method in this paper not only has a better effect of removing raindrops, but also restores the image details to a greater extent.

In order to verify the effectiveness of this method in removing raindrops and improve the recognition capability of computer vision system, yolov3 target detection algorithm was used to test the unprocessed images and the images processed by the method[8].
Figure 6. (a) (c) Yolov3 results of the unprocessed images. (b) (d) Yolov3 results of the processed images.

As shown in figure 6, after yolov3 tests (a) in the car is the train by false identification, and identify the confidence level of 0.94, (b) in the diagram correctly identified the cars, and improve the confidence level of 0.98, (c) in the diagram correctly identified the truck and confidence level is 0.87, (d) improves the recognition degree of confidence in the figure 1. Thus it can be seen that the image processed by the method in this paper can effectively improve the recognition ability of computer vision system.

4. conclusion

This paper proposes a generative adversarial network image removal raindrop method based on difference learning. This method uses the generative network to learn the difference between images with raindrops and clean images. The discriminative network uses matrix output rather than single value output to better solve the problem of generating fuzzy image. Compared with other methods, the method in this paper is superior to other methods both in the visual perception of the image after removing raindrops and in the quantitative test of PSNR and SSIM. Moreover, after testing on the yolov3 target detection algorithm, the image processed by this algorithm can significantly improve the recognition rate of the yolov3 target detection algorithm.

References

[1] P. C. Barnum, S. Narasimhan, and T. Kanade. (2010) Analysis of rain and snow in frequency space. J. Int’l Journal of Computer Vision, 86(2-3):256–274.
[2] Kurihata H, Takahashi T, Ide I, et al. (2005) Rainy weather recognition from in-vehicle camera images for driver assistance. In: Intelligent Vehicles Symposium. Proceedings. IEEE. pp:205–210.
[3] Yamashita A, Tanaka Y, Kaneko T. (2005) Removal of adherent waterdrops from images acquired with stereo camera. In: Intelligent Robots and Systems. IEEE. pp:400-405.
[4] Eigen D, Krishnan D, Fergus R. (2013) Restoring an Image Taken through a Window Covered with Dirt or Rain. In: Proceedings of the 2013 IEEE International Conference on Computer Vision. IEEE. pp:633-640.
[5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al. (2014) Generative adversarial nets. J. Advances in Neural Information Processing Systems:2672-2680
[6] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, et al. (2017) Image-To-Image Translation with Conditional Adversarial Networks. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp:1125-1134
[7] SIMONYAN K, ZISSERMAN A. (2014) Very deep convolutional networks for large-scale image recognition. J. Computer Science, 9 (4):1409-1556.
[8] Redmon J, Farhadi A. (2018) YOLOv3: An incremental improvement. http://arxiv.org/abs/1804.02767v1.