이미지의 눈제거를 위한 심층 Resnet

만위국, 이효종*
전북대학교 컴퓨터공학부
E-mail: wanwgplus@gmail.com, hlee@chonbuk.ac.kr

Deep Residual Networks for Single Image De-snowing

Weiguo Wan[1], Hyo Jong Lee[1, 2, *]
[1] Division of Computer Science and Engineering, Chonbuk National University
[2] Center for Advanced Image and Information Technology, Chonbuk National University

Abstract

Atmospheric particle removal is a challenging task and attacks wide interests in computer vision filed. In this paper, we proposed a single image snow removal framework based on deep residual networks. According to the fact that there are various snow sizes in a snow image, the inception module which consists of different filter kernels was adopted to extract multiple resolution features of the input snow image. Except the traditional mean square error loss, the perceptual loss and total variation loss were employed to generate more clean images. Experimental results on synthetic and realistic snow images indicated that the proposed method achieves superior performance in respect of visual perception and objective evaluation.

1. Introduction

Many computer vision applications are impeded by atmospheric phenomena like rainstorms, haze, snowfall [1]. For example, the bad weather may disturb intelligent surveillance systems and lead to insecure machine interpretation and higher risks of incorrect alarms [2]. Fig. 1 shows an objective detection example with the ImageAI [3]. From it, we can see that the snow particles obstructed the effect of object detection and more objects are detected after removing the snow particles with the proposed method.

In recent years, there are many atmospheric particle removal methods have been put forward to reduce particle obscuration. The common research topics are haze and rain particles removal from images. The haze particle removal methods are designed with a strong assumption, attenuation prior and feature learning, according to the fact that haze particles are uniformly accumulated over an image [4]. The rain particle removal aims at modeling general characteristics of rain images such as shapes, patterns or edge orientations [5]. Nowadays, deep learning-based methods have attracted wide attention due to their better generalization abilities than previous hand-crafted features.

Even though the deep learning-based rain and haze removal approaches are able to locate and remove atmospheric particles, it is difficult to directly apply them in snow removal due to the complex characteristics of snow particles, such as uneven density, diversified particle sizes, irregular trajectory and transparency. Thus, snow removal is more difficult to accomplish and inapplicable for other deep learning-based approaches. The existing snow particle removal methods mainly focus on extracting hand-crafted features, which may lead to weak generalization ability as traditional rain and haze removal methods [6].

In this paper, an end-to-end deep residual network was designed to eliminate the snow particles from single image.

Fig. 1. Object detection on realistic winter snowy image (top) and corresponding snow removal image of the proposed method (bottom) by ImageAI library.

The proposed network was built on U-Net [7] architecture and inception modules. In additional, the perceptual loss and total variation loss were utilized to obtain better performance. Experimental results conducted on both synthetic and realistic snow images demonstrated that our method outperforms the comparison approaches.

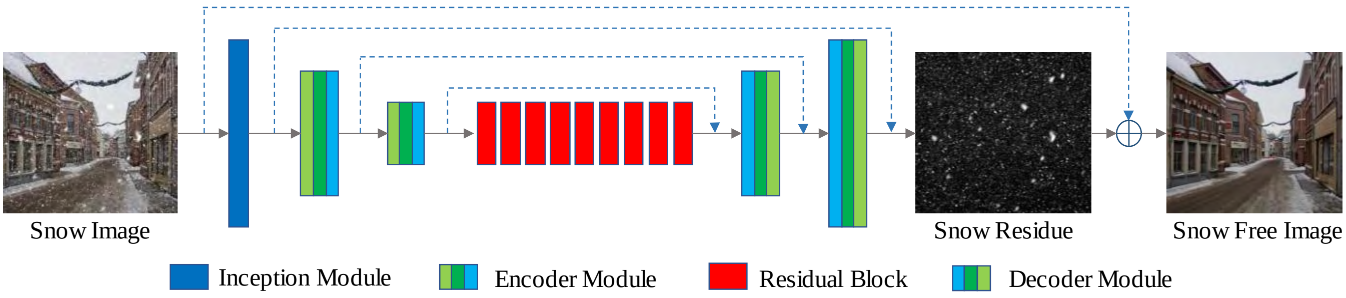

Fig. 2. The overall flowchart of the proposed snow removal method.

2. Deep Residual Networks for Snow Removal

In this section, we will introduce the overall architecture of the designed deep residual networks. As shown in Fig. 2, an inceptual module is first adopted to extract the multiple resolution features of the input snow image, then two convolutional layers with stride 2 are used as encoder to encode the extracted multiple resolution features. Afterwards, nine residual blocks are inserted between the encoder and decoder to increase the learning ability. In the decoder part, two deconvolutional layers with stride 1/2 are utilized to upsample the feature maps to the original resolution. Then, a convolutional layer is employed for converting the feature maps to obtain the target snow residue. Finally, the snow free image can be obtained by adding the snow residue onto the input snow image.

2.1. Inception Module

Due to the fact that the snow image has different sizes of snow particles, thus, we adopted an inception module [8] for multiple resolution feature extraction. The inception module concatenates features generated from three groups of convolutional layers with filter size of 1×1, 3×3, and 5×5 respectively, as shown in Fig. 3. Each group of convolutional layer has 32 filter kernels. Then, a 3×3 convolutional layer with 64 filters is operated on the concatenated feature maps to obtain the output of the inception module.

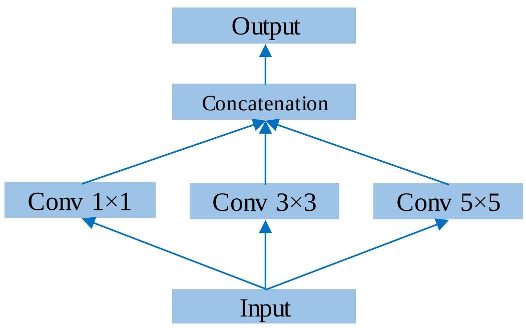

Fig. 3. The illustration of inception module.

2.2. Residual Blocks

Recently, the residual networks have shown excellent performance in computer vision tasks, e.g. image super-resolution [9], image deblur [10], raindrop removal [11] et al. In a residual network, each layer is fed into the next layer and directly into several hops away, as shown in Fig. 4. The residual networks can overcome vanishing gradient problem when calculating backpropagation through a very deep network. In this paper, we employed multiple residual blocks to increase the learning ability. Different from the original residual module, we optimized it by replacing the batch normalization with the instance normalization which has been demonstrated can obtain better results in image-to-image generation tasks [12].

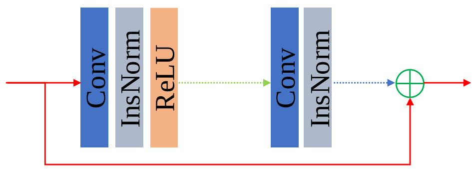

Fig. 4. The architecture of a residual block.

2.3. Loss Function

The appropriate loss functions are crucial to better train network models. Mean squared error (MSE) or L2 loss is the most widely used loss function for general image restoration tasks. Assume that y and y' are the snowy image and the corresponding snow free image respectively, n is total number of pixels in each image. The MSE loss can respected as follow:

$$L_M(y, y') = \frac{1}{n}\sum (y - y')^2, \tag{1}$$

Recently, the perceptual loss [13] has been improved to be an effective way to deal with the style by extracting the high-level features from deep networks, like VGG-19. The perceptual loss can be expressed as:

$$L_p(y, y') = \sum \| \Phi_l(y) - \Phi_l(y') \|_2^2, \tag{2}$$

where Φ_l is the l-th feature map of the VGG-19 network.

In addition, we utilize the total variation loss to eliminate noise and artifact of the predicted snow free image:

$$L_T(y') = \sum_{m,n}((y'_{i+1,j} - y'_{i,j})^2 + (y'_{i,j+1} - y'_{i,j})^2), \tag{3}$$

where $y'_{i,j}$ is the pixel value at (i, j) of the predicted snow free image.

The total loss function is calculated by weighted averaging the above loss functions:

$$L = \lambda_1 L_M + \lambda_2 L_P + \lambda_3 L_T. \tag{4}$$

3. Experiments

3.1. Dataset

In this paper, the Snow 100K dataset [14] was used to train and test the proposed deep snow removal network. There are 100k synthetic snowy images, 100k corresponding

(a) (b) (c) (d)

Fig. 5. Snow removal results of various methods for synthetic snowy images. (a) Input image. (b) Zheng et al. [15]. (c) DeblurGAN [10]. (d) The proposed method. Best viewed on screen.

snow free ground truth image, and 1329 realistic snowy images in this dataset. All the images were normalized with the size of the largest boundary of each image to 640 pixels and retained their original aspect ratio. We randomly selected 10k image-pairs from the synthetic dataset as training data, and randomly selected 2k image-pairs from the rest as testing data. All the realistic snowy images were used for testing.

3.2. Implementation details

In this section, the implementation details about the proposed method and experiment are introduced. The λ values in (4) are [1, 10, 0.0001]. The training epochs is 10, and the batch size is 8. The Adam optimizer with learning rate 0.0001 is employed for all the networks training. Our experiments were performed on a computer with Intel(R) i7, 8GB RAM and NVIDIA GTX Titan-X GPU.

3.3. Experimental Results

To demonstrate the performance of the proposed method, one traditional method [15] and one deep learning-based method [10] were compared. For fair comparison, we retrained the deep learning model with snow training images.

The experiments were conducted both on synthetic and realistic snow image datasets. The source snow images and the predicted snow free images by different methods are displayed in Fig. 5 and Fig. 6. From the visual perception comparison, we can see that the method of Zhang et al. cannot eliminate the snow particles very well and results in

(a) (b) (c) (d)

Fig. 6. Snow removal results of various methods for realistic snowy images. (a) Input image. (b) Zheng et al. [15]. (c) DeblurGAN [10]. (d) The proposed method. Best viewed on screen.

serious blur effect. The result images of the DeblurGAN method exist many artifacts. Our method obtains best results in eliminating snow particles and preserving image details. The evaluation indexes on PSNR and SSIM also indicated the superiority of the proposed method, as shown in Table 1.

Table 1. Objective comparison of different methods.

	Synthetic	Zheng et al.	DeblurGAN	Proposed
PSNR	22.1793	22.3247	23.3019	**24.7814**
SSIM	0.8254	0.7532	0.8273	**0.8549**

4. Conclusions

This paper presented an end-to-end image single snow removal approach based on deep residual network. The inception module was utilized to extract multiple resolution features. Moreover, perceptual loss and total variation loss were adopted to train the designed network. The synthetic and realistic snow images were used to evaluate the performance of the proposed method, and the experimental results indicated the superior effectiveness of the proposed method. In future works, we plan to develop better deep networks and apply our method in other atmospheric particle removal tasks.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (GR2016R1D1 A3B03931911).

References

[1] Deng L., Huang T., Zhao X., et al. A directional global sparse model for single image rain removal. Applied Mathematical Modelling, 2018, 59: 662-679.

[2] Fu X, Liang B, Huang Y, et al. Lightweight pyramid networks for image deraining. arXiv preprint arXiv:1805.06173, 2018.

[3] http://imageai.org/.

[4] Li R, Pan J, Li Z, et al. Single image dehazing via conditional generative adversarial network. in Proceedings of the CVPR. 2018: 8202-8211.

[5] Yang W, Tan R T, Feng J, et al. Deep joint rain detection and removal from a single image. in Proceedings of the CVPR. 2017: 1357-1366.

[6] Pei S, Tsai Y, Lee C. Removing rain and snow in a single image using saturation and visibility features. IEEE International Conference on Multimedia and Expo Workshops. 2014: 1-6.

[7] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. 2015: 234-241.

[8] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.

[9] Zhang Y, Tian Y, Kong Y, et al. Residual dense network for image super-resolution. in Proceedings of the CVPR. 2018: 2472-2481.

[10] Kupyn O, Budzan V, Mykhailych M, et al. Deblurgan: Blind motion deblurring using conditional adversarial networks. in Proceedings of CVPR. 2018: 8183-8192.

[11] Fan Z, Wu H, Fu X, et al. Residual-guide network for single image deraining. 2018 ACM Multimedia Conference on Multimedia. 2018: 1751-1759.

[12] Dmitry U, Andrea V, and Victor L. Improved texture networks: Maximizing quality and diversity in feed-forward stylization and texture synthesis. In CVPR, 2017.

[13] Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. European conference on computer vision. 2016: 694-711.

[14] Liu Y F, Jaw D W, Huang S C, et al. DesnowNet: Context-aware deep network for snow removal. IEEE Transactions on Image Processing, 2018, 27(6): 3064-3073.

[15] X. Zheng, Y. Liao, W. Guo, X. Fu, and X. Ding, “Single-image-based rain and snow removal using multi-guided filter,” in Proc. Int. Conf. Neural Inf. Process., pp. 258–265, 2013.