Research Article

Single Image Defogging Algorithm Based on Conditional Generative Adversarial Network

Rui-Qiang Ma,^{1,2} Xing-Run Shen,² and Shan-Jun Zhang²

¹Inner Mongolia University of Technology, Huhhot 010080, Inner Mongolia, China

²Kanagawa University, Hiratsuka, Kanagawa 259-1293, Japan

Correspondence should be addressed to Shan-Jun Zhang; zhang@info.kanagawa-u.ac.jp

Received 13 October 2019; Revised 6 May 2020; Accepted 20 September 2020; Published 24 November 2020

Academic Editor: Javier Martinez Torres

Copyright © 2020 Rui-Qiang Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Outside the house, images taken using a phone in foggy weather are not suitable for automation due to low contrast. Usually, it is revised in the dark channel prior (DCP) method (K. He et al. 2009), but the non-sky bright area exists due to mistakes in the removal. In this paper, we propose an algorithm, defog-based generative adversarial network (DbGAN). We use generative adversarial network (GAN) for training and embed target map (TM) in the anti-network generator, only the part of bright area layer of image, in local attention model image training and testing in deep learning, and the effective processing of the wrong removal part is achieved, thus better restoring the defog image. Then, the DCP method obtains a good defog visual effect, and the evaluation index peak signal-to-noise ratio (PSNR) is used to make a judgment; the simulation result is consistent with the visual effect. We proved the DbGAN is a practical import of target map in the GAN. The algorithm is used defogging in the highlighted area is well realized, which makes up for the shortcomings of the DCP algorithm.

1. Introduction

Fog is a natural phenomenon that blurs the scene, reduces visibility, and changes color. This is an annoying problem for photographers because it can reduce image quality. It also threatens the reliability of many applications, such as outdoor surveillance, object detection, and aerial imaging. Therefore, in the computer vision or graphics technology, it is very important to remove the blurs in the image. However, due to its mathematical fuzziness, usually only an image is input, and it is very challenging to eliminate the fuzziness.

Considering the particularity of fog image, image processing is based on prominent details, contrast enhancement, and brightness enhancement, which has a certain image enhancement effect in vision, and it is not the result of defogging in essence. In histogram equalization, by redistributing image pixel values, the number of pixels in a certain gray range is approximately the same [1]. This method enhances the global contrast of the image and enlarges the image noise easily. In order to overcome the shortcomings, Zuiden's group proposed a contrast-limited adaptive

histogram equalization algorithm [2], which limits the magnification by clipping the histogram with a predefined threshold before calculating the cumulative distribution function, but it is easy to make the color image color skewed. In terms of image restoration methods based on physical models, some researchers use the atmospheric scattering model [3] proposed by Mcica's group to model hazy scenes to solve the problem of hazy degraded images. For example, Tan's group and other scholars have found that fog-free images have higher contrast than fog images [4], which maximizes the local contrast of images, but the restored color is often oversaturated. Specifically, the atmospheric scattering model is based on prior knowledge: one is the fog-free images that have higher contrast than images under severe weather conditions, and the other is light that is often smooth between the scene and the observer. Then, local pairing of scene images is improved by using the Markov random fields (MRF) model. Ratio restored the fog-free image of the scene. The result of this processing method makes the contrast of the image greatly improved, but there is a large halo phenomenon, and the scene appears with a

certain degree of color deviation, which leads to the unreal image.

In order to solve the problem of excessive distortion of defogging images, in 2009 HeKm's team of the Chinese University of Hong Kong proposed a single image defogging technology, that is, dark channel prior algorithm [5], and won the CVPR Best Paper Award of that year, which made the image defogging technology a landmark breakthrough. Based on the statistical characteristics of outdoor daytime fog-free images, the brightness value of pictures in at least one color channel of R, G, and B is close to zero, and the dark channel prior theory was proposed. Combining this prior knowledge with the atmospheric scattering model, the transmittance is obtained by the dark channel prior algorithm and then corrected by the soft matting algorithm to estimate atmospheric light value. Finally, a better restoration of the fog image is achieved. The dark channel prior method has made a new leap in the field of image defogging, and many subsequent research methods are mostly based on this method.

The principle of DCP belongs to a color image $J(x)$ with R, G, and B channels, and two minimum filters are performed on R, G, and B channels and filter template $\Omega(x)$ of $N \times N$ size. Normally, N is set to 15, and through statistics of a large number of outdoor fog-free images, it is concluded that the dark channel of outdoor fog-free images has properties.

To get atmospheric light value A , select the brightest 0.1% pixels in the dark channel image $I(x)$, where the fog is thickest and then correspond these points to the same position of the original image $I(x)$. In these points of the original image $I(x)$, select the maximum value of three channels as the atmospheric light value, respectively. Ray propagation map is $t(x)$: calculate dark channels on both sides of the equation at the same time.

1.1. Hazy Image Formation Model. The hazy image degradation model consists of two parts: the scene reflection light attenuation model and the atmospheric light imaging model, which is proposed by McCartney [6]. Nayar and Narasimhan [7] further derived the model later, and the mathematical models used to describe the above processes are given in equations (1) and (2), as follows:

$$I(x) = J(x)t(x) + A(1 - t(x)), \quad (1)$$

where $I(x)$ is the light intensity received by the observation point, $J(x)$ is the light intensity radiated by the scene point, that is, the haze-less image to be restored, $t(x)$ is the light transmission map, $J(x)t(x)$ is the scene reflection (or emission) light attenuation model, $A(1 - t(x))$ is the atmospheric light value, and $A(1 - t(x))$ is the atmospheric light imaging model.

$$t(x) = e^{-\beta d(x)}, \quad (2)$$

where $d(x)$ is the distance between the observation point and the scene point, β is the atmospheric light scattering coefficient, and β is regarded as a constant in the visible light range [8].

With the dark channel prior (DCP) algorithm, it is possible to achieve the ideal implementation of dehaze treatment but also occurred in the bright area of the miss restore, as shown in Figure 1, in which Figure 1(c) is the local magnification of Figure 1(a), and Figure 1(d) is the local magnification of Figure 1(b).

In recent years, researchers have proposed many prior hypotheses to obtain some statistical or decisive properties of fog maps. Tufail's group proposed transmission maps to reconstruct the images with different color contrast [9] and the DCP-based image defogging method with improved transmission map to avoid blocking artifacts. The transmission maps are computed for RGB and YCbCr color spaces [7]. A prior defogging algorithm has achieved good results on some fog maps. These algorithms require fog maps to fully conform to scenario assumptions. Although a prior defogging algorithm achieves defogging, this kind of algorithm has the disadvantage of weak applicability. In the real world, many fog maps do not always conform to the hypothesis, so the results of fog removal of these algorithms will have a variety of problems. Several typical cases are as follows. For example, Fattal's algorithm is suitable for fog with light fog, but it is prone to noise in fog with heavy fog. Because of its rich color information and the need for color difference between pixels, these requirements are not available in fog maps. He's algorithm is not suitable for sky or bright scenes because He's method assumes that atmospheric light intensity is brighter than all pixels. However, in a picture with a bright sky or light source, there will be some brighter points than the light. He's method estimates the transmittance of these pixels very small, even in some negative numbers. In fact, the normal range of transmittance should be between 0 and 1. After the final restoration, the color difference appears in the sky region of the image. Berman's method is not suitable for images whose light intensity is obviously brighter than other pixels.

2. Our Method

In the dark channel prior method, when the image is similar to airlight over a large local region and no shadow is cast on the image of object, conditional GAN is used to carry out for confrontation training, and then the problem is solved. In addition, much of the information in the fog-less image has been lost. Finally, according to the convergence of discriminator loss and generator loss [10], the training end time point is determined, and the training model is obtained as the basis for processing the unprepared area of the bright area of the image again (Figure 2).

The innovation of this work is to use the RGB information data after defogging of the DCP method and use the RGB information data of the fog-less image as the counter training; it has solved the unavoidable confusion of defogging distortion of the bright part of DCP effectively [11, 12].

In order to solve the problem of bright distortion caused by the DCP method, the generation network and output discrimination network are used to evaluate it so as to ensure that the output looks like the fog-less image. In this process,

FIGURE 1: Miss restored image. (a) Hazy image, (b) DCP algorithm, (c) local enlargement of (a), and (d) local enlargement of (b).

FIGURE 2: Image of DbGAN composition.

the target map is generated firstly. Target map is an important part of the network because it will guide the network to focus on the haze distortion area, and target map is generated by the cyclic network.

Then, the generated network uses the designed automatic encoder to input the RGB information of the image as the input, referring to the target map. In order to obtain more extensive context information, multiscale loss is used in the decoder side of the automatic encoder. Each loss compares the output of the convolution layer with the corresponding ground truth. The input of the convolution layer contains the characteristics of the decoder layer. In addition to get the data of these losses, for the final output of the automatic encoder, a perceptual loss is used to obtain a more comprehensive similarity with the ground truth. The final output is the output of the generated network too. In formula (3), $f(x)$ and $g(x)$ are the feature vectors of the front layer, and N is the number of front hidden layer features. The model that synthesizes the No. j area integrates the degree of participation at the No. i position, and local information and all stations are integrated.

$$o_j = v \left(\sum_{i=1}^N \left(\frac{\exp(f(x_i)^T g(x_j))}{\sum_{i=1}^N \exp(f(x_i)^T g(x_j))} h(x_i) \right) \right). \quad (3)$$

Through the training, the target layer information can be obtained as output; furthermore, by adjusting the super parameters $\gamma(0, \infty)$ can be obtained with scalar, and TM can be put into the GAN model gradually, as follows:

$$y_i = \partial o_j + x_i. \quad (4)$$

After obtaining the output of the generated image, the discrimination network will check whether the value is true or not. In this study, target map is used to guide the discrimination network to the local target area. In general, target map is introduced into the generation network and discrimination network, which is a new method. It can effectively achieve haze removal which is impossible for feature extraction and defogging DCP method.

The algorithm of DbGAN is given as Algorithm 1, $p_g(z)$ is the distribution of fogging image sets, and $p_{data}(x)$ is the vector of fog-less image sets.

It must update the discriminator and generator for $p_g(z)$ approaches $p_{data}(x)$ and $P_G(x; \theta)$, where θ is the parameter.

2.1. Experimental Results. In the generation countermeasure network, there are generation network and discrimination network. Receiving and inputting an image information with incomplete processing, the generating network tries to generate a bright area for processing as much as possible to achieve the effect of bright area defogging. The discrimination network will verify whether the image generated by

```

(1)  for (number of training iterations){
(2)    for (hyperparameter){
(3)      minibatch fogging images  $\{z^{(m)}\}$  with  $p_g(z)$ ;
(4)       $T(m) = f_m(C_{m-1}, T(m-1))$ ; //introducing the target map
(5)      minibatch fog-less image  $x^{(m)}$  with  $p_{data}(x)$ ;
(6)      update the discriminator;
(7)    }
(8)    sample minibatch of  $m$  hazy samples  $p_g(x)$  with  $p_g(z)$ ;
(9)    update the generator;
(10)  }

```

ALGORITHM 1: The algorithm of DbGAN.

the generated network looks real. Its loss can be expressed as follows:

$$\min_G \max_D E_{R \sim P_{\text{cicn}}} [\log(D(R))] + E_{I \sim P_{\text{raindrop}}} [\log(1 - D(G(I)))]. \quad (5)$$

This kind of haze removal algorithm is based on RGB after the DCP algorithm. Using the generated counter-measure network, the generated network generates the target map through the attention recurrent network and generates the highlighted area restored image through the contextual autoencoder together with the input image. Then, the global and local validity of the output is generated by judging the network evaluation. In order to be able to verify locally, we inject target map into the network, which is also the innovation of this method, and we use target map in generating network and discriminating network. For the fog-less image and fogging image of the same object, it is hoped that the output of this image with haze will approach the fog-less image after the network processing infinitely.

The model was trained in this work, and the data fitting of the results showed that the accuracy of training and validation data tends to be 1 and then tends to be stable in a certain period of time rather than rising. Therefore, the approximate value of the epoch can be determined.

2.2. Target Map and RGB Values. The new algorithm introduced the target map to the GAN method, and it is a model that mimics the neural circuits of the human brain. The visual map of gaze areas can be obtained by visual interpretation, and the receptive field of each layer of target characteristics can be improved. It amounts to increase the depth of the network. For example, as shown in Figure 3, in the original image, areas marked with red, yellow, and blue are considered as sensitive parts of human eyes. Through the TM model, we can notice these areas and realize centralized processing, which improves efficiency. As shown in Figure 3, the visible area in the original image will be displayed in the layer. These conspicuous areas are often the location of the DCP miss restored area [4].

In these cases, the convergence is ideal and the stable correlation values of train accuracy and test accuracy tend to be 1. The convergence is obvious, and there are local small

fluctuations in them. However, the use of the generated model has little impact on the final processing results. Three examples are used for further illustration, as shown in Figure 4

According to the model obtained from training, the demos obtained from restoration are shown in the following example, and the corresponding RGB three values, as well as PSNR, are shown in Figure 4 from demo1 to demo3.

The RGB value of the defogged image of DCP is calculated after the defogged image is generated by DCP. In the DbGAN algorithm and images training, we can obtain defogging images and their RGB values calculated. Using this kind of GAN, we can obtain fog-less image to export an ideal model. Finally, we calculated the RGB value of the restored image. As shown in the figure, the last line of the three demos is the RGB images obtained through the DbGAN confrontation training.

The PSNR test value in the demos shows that, for the demo case, three different defogging algorithms are used, respectively. According to the order of DCP and DbGAN, the corresponding PSNR value increases in turn, indicating that the image reconstruction quality is better.

In these three demo cases, the white light area has different degrees of false repair, or exaggerated restoration, and the distortion component is relatively large. After DbGAN processing, the haze removal effect of the white light area has been fully realized.

2.3. Simulation and Conclusions. In this research, the main equipment configuration parameters for training are as follows.

There are Intel Core i5-7300HQ CPU @ 2.5 GHz and NVIDIA GeForce GTX 1050 Ti. Then, software include OpenCV4.01 + Visual Studio2011 (Win10), C/C++ & Python, and Anaconda3 (Jupyter and PyCharm).

The database was used, and then three sets of data systems were collected and downloaded to form the data set for our research, which are CIFAR-10, ImageNet, and benching mark [13].

In order to verify the effectiveness of the proposed dehaze method, we validated it on various hazy images and compared with old methods.

MSE represents the mean square error of image I and image J ; the lower the value, the better the quality.

FIGURE 3: Visualization of target map.

FIGURE 4: RGB images generated by DbGAN. (a) Demo 1. (b) Demo 2. (c) Demo 3.

TABLE 1: Comparison of PSNR values of the image in Figure 5.

No.	PSNR (dB)	
	DCP	DbGAN
a	9.93	10.58
b	12.74	17.79
c	14.35	17.76
d	11.57	13.28
e	10.85	17.50
f	9.19	10.70
g	8.46	13.27
h	9.56	13.27
Aver	10.83	15.47

$$\text{MSE}(X, Y) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2. \quad (6)$$

Then, using MSE, the definition of solving PSNR is as per formula (7), and it can be obtained as follows:

$$\text{PSNR}(X, Y) = 10 \lg \left(\frac{\text{MAX}^2}{\text{MSE}} \right). \quad (7)$$

For the evaluation, the different evaluation indexes are adopted due to different objectives. MSE is usually used, but in our research, we used the PSNR evaluation method. The PSNR test results are given in Table 1. However, MSE prepared a correct answer picture and compared each picture element, respectively, in formula (6). About MSE, the smaller value is good, and the image that is closer to the original image is mainly used for the evaluation of restoration processes. MAX is the biggest picture element value

that a former picture can take. In this way, a clear image is obtained while maintaining the haze removing quality.

Comparison of research proposals is shown in Figures 5 and 6.

From the data shown in Table 1, it can be concluded that although the DCP method has high reconstruction quality in theory, the DbGAN method is more consistent with the human visual characteristics, while the PSNR value of the DbGAN algorithm is basically close.

We have investigated mountains, fields, woods, people, Tian'anmen Square, buildings, goose, train lights, and all eight typical haze patterns; we compared the processing results of two defogging algorithms, DCP and DbGAN. Simply from the PSNR values of each algorithm, except for fields and woods, the values of other images increase in turn, which shows that the image quality reconstruction is getting better and better. The field image basically does not have the problem of special repair of the highlighted area. The fogging image is relatively uniform,

FIGURE 5: Comparison of three algorithms.

the undulation is not obvious, and the background is simple, so DCP or DbGAN can well correspond.

In the field of digital image processing, the commonly used evaluation indexes include PSNR and MSE. Two categories will be used in this research as peak signal-to-noise ratio; the PSNR values before and after image restoration indicate the ratio of the maximum power of a signal. Generally, the higher the PSNR value, the better the image reconstruction quality. The visual characteristics of human eyes are not considered. Because the sensitivity of the human eye to the contrast difference with lower spatial frequency is higher and the sensitivity of the human eye to the contrast

FIGURE 6: Details of images.

TABLE 2: Comparison of PSNR values of the details in Figure 6.

No.	Light image PSNR (dB)	Goose image PSNR (dB)
DCP	8.77	8.64
DbGAN	8.51	14.31

difference of brightness is higher than chroma, the perception result of the human eye to an area will be affected by the adjacent area around it, so the evaluation result is often different from the subjective feeling of the human.

Although the data that affect images do not make difference in bright areas, the difference is large in nonbright areas, and their values increase (Table 2), which indicates that the processing effect is better (Figure 7). This index shows that the larger the value, the greater the similarity of the image.

The effect of defogging is more obvious from the enlarged drawing and the detailed inspection; the goose image is closer to the original image than the train lights because the light area of the train light area is larger than that of the goose area. In the word image, the evaluation parameters are the results of the research study conducted by the auxiliary researchers. On this basis, it is necessary to cooperate with the visual inspection results to achieve better results. For example, the results of train lights for two different algorithms differ greatly after zooming in, but people can feel very sensitively that the effect of the DbGAN should be the closest to the reality.

FIGURE 7: Values of PSNR.

However, the human eye has some differences compared with the evaluation index, for example, high sensitivity to low spatial frequency contrast. The difference in brightness is more sensitive than the difference in saturation. Perception result for a certain area is affected by surrounding information.

3. Conclusions

For single image defogging, the DbGAN algorithm was proved that it solved the wrong removal processing in the DCP method; the bright areas of the image has been removed obviously. Using target map, the content in the conditional GAN network architectures was submitted. Further, as the target map, the feature map extraction function of the geometric architecture of the image is demonstrated; for each layer, the receptive field of the feature is spread. It is possible to output a focused layer by training while being a good complement to define the features simply of centering on the GAN texture. As a result, the effect of defogging in the highlighted area is well realized, and then it makes up for the shortcomings of the DCP algorithm. Subsequently, in the PSNR evaluation indicator, the advantages are obvious in image quality and calculation efficiency. The higher the PSNR values, the better the removal of distortion. In our verification experiment, the average value of PSNR is higher than DCP, consistent with the results of visual verification.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was funded by the Innovation and Entrepreneurship Foundation of Overseas Students in Inner Mongolia (2019), the Chunhui Plan of the Ministry of Education (2019), the Inner Mongolia Science and Technology Project (2019GG372), the Inner Mongolia Natural Science Foundation Project (2020MS06025), and the NSFC under Grant No. 61762070.

References

- [1] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 2995–3002, Columbus, OH, USA, June 2014.
- [2] J. Kim, J. Lee, and K. Lee, "Accurate image super resolution using very deep convolutional networks," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1646–1654, Las Vegas, NV, USA, June 2016.
- [3] C. Liu, J. Zhao, Y. Shen, Y. Zhou, X. Wang, and Y. Ouyang, "Texture filtering based physically plausible image dehazing," *The Visual Computer*, vol. 32, no. 6–8, pp. 911–920, 2016.
- [4] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, June 2009.

- [5] G. Cui, H. Feng, Z. Xu, Q. Li, and Y. Chen, "No-reference image noise estimation based on noise level accumulation," *Optical Review*, vol. 23, no. 2, pp. 208–219, 2016.
- [6] E. J. McCartney, *Optics of the Atmosphere, Scattering by Molecules and Particles*, Wiley, New York, NY, USA, 1976.
- [7] Z. Khurshid, K. Khurshid, A. Salman, I. Fareed Nizami, K. Khurshid, and B. Jeon, "Improved dark channel prior for image defogging using RGB and YCbCr color space," *IEEE Access*, vol. 6, no. 1, pp. 32576–32587, 2018.
- [8] Y. Shen, C. Liu, J. Zhang, Y. Shao, and J. Zhao, "Atmospheric light correction and transmission optimization based robust image dehazing," *Journal of Image and Graphics*, vol. 29, no. 9, pp. 1604–1612, 2019.
- [9] Z. Tufail, K. Khurshid, A. Salman, and K. Khurshid, "Optimisation of transmission map for improved image defogging," *IET Image Processing*, vol. 13, no. 7, pp. 1161–1169, 2019.
- [10] H. Zhang and V. M. Patel, "Density-aware single image deraining using a multi stream dense network," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 695–704, Salt Lake City, UT, USA, June 2018.
- [11] Y. Li, N. Xiao, and W. Ouyang, "Improved boundary equilibrium generative adversarial networks," *IEEE Access*, vol. 6, pp. 11342–11348, 2018.
- [12] Q. Yang, P. Yan, Y. Zhang et al., "Low dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss," *IEEE Transactions on Medical Imaging*, IEEE, vol. 37, no. 6, pp. 1348–1357, 2018.
- [13] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017.