The Improved Pix2pix Generative Adversarial Networks for Sand-dust Image Enhancement

Zhongwei Hua  
Fudan University

Lizhe Qi (qilizhe@fudan.edu.cn)  
Fudan University

Zhi Yang  
Fudan University

Yunquan Sun  
Fudan University

Research Article

Keywords: sand-dust, image enhancement, generative adversarial networks

Posted Date: October 29th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2191083/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
The Improved Pix2pix Generative Adversarial Networks for Sand-dust Image Enhancement

Zhongwei Hua, Lizhe Qi*, Zhi Yang and Yunquan Sun

Academy of Engineering and Applied Technology, Fudan University, Handan road 220, Shanghai, 200433, Shanghai, China.

*Corresponding author(s). E-mail(s): qilizhe@fudan.edu.cn;

Abstract

The frequent sand-dust weather in inland areas has severely affected the local outdoor computer vision applications. Different from the previous ideas of the sand-dust image enhancement algorithm, this paper proposes a generative adversarial network to enhance the sand-dust images. We improve the classic pix2pix network by introducing the dual attention mechanism to the U-net architecture and improve the loss function of the generator through Smooth L1 and SSIM to further enhance the color reproduction, detail features, the structural similarity, and the convergence speed of the generator. In addition, we also publish the first artificially synthesized sand-dust image data set online. The experimental results show that the enhancement method proposed in this paper has obvious advantages in both artificially synthesized images and natural real images, compared with the current traditional sand-dust enhancement algorithms and the previous network models.

Keywords: sand-dust, image enhancement, generative adversarial networks

1 Introduction

At present, the area of global land desertification is more than 33% of the earth’s land[1], and the resulting sand-dust weather is increasing day by day. Severe sand-dust weather not only harms human health, but also severely affects the effects of outdoor computer vision applications such as self-driving, highway monitoring, and target detection. In sand-dust weather, there are a large number of dust particles floating in the air, the absorption and scattering of light are very serious, and the image will produce obvious color distortion, contrast degradation and blurred details. Therefore, it is necessary to eliminate the impact of sand-dust weather on the image.

Different from traditional image enhancement methods, deep learning-based methods use data-driven methods and recently have achieved significant results in image enhancement during severe weather. Among them, the strong creation and transformation capabilities of the Generative Adversarial Network (GAN) has attracted widespread attention from scholars. The pix2pix[2] model composed of a U-net structure generator and a PatchGAN structure discriminator is one of the most popular conditional-based generative adversarial network models. The network model requires pairs of original images and target images as training data. Through the iterative game between the discriminator and the generator, the generator finally has excellent image conversion capabilities.

Therefore, we propose an improved pix2pix network to enhance the sand-dust images, adding...
the dual attention mechanism (channel attention and spatial attention) to the U-net architecture, and improving the loss function of the generator through Smooth L1 and SSIM. Experimental results show that this method has advantages in enhancing both artificial synthetic images and natural real images.

2 Related Work

Based on traditional image processing ideas, many scholars have proposed enhancement algorithms specifically for sand-dust images. Fu et al. [3] first proposed a special enhancement algorithm for sand-dust images, used a weighted fusion of two gamma-corrected images with three color details. Li et al. [4] mapped the sand-dust image to the Lab color space, achieved color correction by compensating and stretching the a and b channels, and sharpened and filtered the L channel to enhance the detailed. Inspired by underwater image enhancement, some scholars [5–8] compensated the blue channel based on the green channel information to correct the color deviation of the sand-dust image. In order to further enhance the contrast of the sand-dust image, Shi et al. [9] proposed a limited contrast adaptive histogram equalization algorithm based on the normalized gamma transform for the L channel. Chen et al. [10] enhanced the detail of sand-dust image with multi-scale exposure fusion. Wang et al. [5] implemented a multi-scale fusion algorithm after color compensation to enhance the sand-dust images. Gao et al. [11] proposed a method of inverting the blue channel before applying the dark channel prior method to effectively solve the problem of sky color distortion after enhancement. Although based on traditional image processing methods, especially in contrast enhancement, it has a good effect on sand-dust image enhancement, most of the enhanced images still have problems such as halos, blue or red artifacts, time-consuming, poor saturation, and low brightness.

Based on deep learning based methods recently have achieved significant results in image enhancement during severe weather. Wu et al. [12] proposed a dehazing network with autoencoder and contrastive regularization (AECR-Net). Chen et al. [13] proposed a Physical Priors Guided Dehazing (PSD) framework to improve the generalization performance of dehazing. Gu et al. [14] proposed a selective predictor (SSEP) to forecast the concentration of PM_{2.5}. Ren et al. [15] proposed a progressive ResNet for rain removal. Although the Generative Adversarial Network has also derived a series of image enhancement methods under severe weather, such as Cycle-Dehaze [16] for haze image enhancement, MBA-RainGAN [17], for rain image enhancement and Enlighten-GAN [18], for low-light image enhancement, there is currently no model for sand-dust image enhancement.

As we all know, it is difficult to collect training samples based on deep learning methods, especially in real sand-dust environments, which may lead to the main reason why deep learning methods have not been applied to the sand-dust image enhancement. Even the artificially synthesized sand-dust image data set has not been produced by researchers or published on the Internet currently, however, the artificially synthesized image data set based on the principle of sand-dust image formation is relatively feasible. Therefore, the main contributions of this paper are:

- We are the first to publicly share the artificially synthesized sand-dust image data set [19] online composed of 400 pairs of sand-dust images and corresponding clear images.
- We propose an improved pix2pix network to enhance the sand-dust images, adding the dual attention mechanism (channel attention and spatial attention) to the U-net architecture, and improving the loss function of the generator through Smooth L1 and SSIM.

3 The proposed method

3.1 The sand-dust image data set

Considering the similarity with haze images, we combined the sand-dust image forming principle to create an artificially synthesized sand-dust image data set based on the public haze data set O-HAZE [20] which contains 45 different outdoor scenes. The idea of generating sand-dust images from haze images will greatly reduce the complexity of creating artificially synthesized sand-dust image datasets.
The impact of severe sand-dust weather on images is far greater than haze. A large number of sand-dust particles are floating in the air. The absorption and scattering of light are very serious. It is found through experiments that the blue channel is most affected by absorption, and the red channel is most affected by scattering. Since the mixed light of green and red is yellow, the sand-dust image often exhibits the yellow or orange color shift. And as shown in Fig. 1, in the wake of the gap between red and green keeps increasing, the color shift is more towards orange. In addition, the contrast and details of the image will appear blurred. The histogram of the sand-dust image shows that the noise is approximately Gaussian.

Finally, considering the low visibility and blurred details of sand-dust images, we blur the adjusted images with different degrees of Gaussian blur and stitch them with the corresponding clear image to form the pair of sand-dust image data sets, as shown in Fig. 2.

### 3.2 Network structure

Considering the powerful image conversion capability of the pix2pix model, we first tried to apply the model to sand-dust image enhancement. We use the artificially synthesized sand-dust image as the original image, and the corresponding natural clear image as the target image and send them to the model training at the same time. In the experiment, it is found that although the pix2pix model has good color cast correction ability, the details of the generated image are seriously lost, and even the main target in the foreground is blurred.

Therefore, we combine channel attention and spatial attention, which can enhance the strongly correlated features and suppress the weakly correlated features by calculating the channel’s global correlation and feature weight into the pix2pix generator structure to focus and enhance the local detailed feature information of the main target in the foreground. We refer to the chain structure of the classic dual attention module CBAM[21], and Many experiments have also verified that the best order of attention is channel attention before spatial attention. The difference is that in the spatial attention part, we use 3*3 dilated convolution to aggregate spatial features of spatial attention, which further expands the receptive field and improves the performance of spatial attention.

As shown in Fig. 3, the generator in this paper adopts a symmetrical U-Net structure of 8 downsamples and 8 up-samples. Each sampling layer includes the fully connected convolution layer, normalization layer and activation function. The convolution kernel is 4 * 4 and the step size is 2. Leaky ReLU activation function is used under each down-samples layer, the ReLU activation function is used for the up-sampling of the first 7 layers, and the Tanh activation function is used in the last layer. The down-sampling convolution layer of each layer is skip connected with the up-sampling convolution layer located in the same layer. At each sampling layer of the model (except for the last up-sampling layer) add the

\[
\begin{align*}
R' &= (1.5 \sim 1.8) \times R \\
G' &= (0.9 \sim 1.3) \times G \\
B' &= (0.2 \sim 0.5) \times B
\end{align*}
\]
dual attention module, that is, the input features are respectively maximized and average pooled, and then the two features are added through the shared multi-layer perceptron, and then the channel attention is generated through the sigmoid activation function, multiplied with the corresponding elements of the input features one by one. The spatial attention part is to generate spatial feature descriptors through the maximum pooling and average pooling operations in series, and uses 3*3 dilated convolution to map the information in the space that needs to be emphasized or suppressed to more efficiently aggregate spatial context information, and then the spatial attention feature and input are generated by the sigmoid function, and then the corresponding elements are multiplied. Finally, the spatial feature and channel feature of the sand-dust image are combined to achieve dual-channel attention to enhance the local detailed feature information of the foreground.

The discriminator in this paper inherits the clever PatchGAN structure in the pix2pix model, that is, the judge will output a judgment matrix, and each value on the matrix corresponds to the receptive field of the original image. Through the segmentation of the original image, the detailed texture information is guaranteed. According to the test, when the Patch is set to 70*70, the image details and colors have better results.

3.3 Loss function

The loss function in the pix2pix model uses a combination of L1 and CGAN. However, it was found in experiments that due to the constant gradient of L1, the training is difficult to converge to higher accuracy, and the color saturation of the enhanced image is still quite different from the original image. This paper uses Smooth L1, which is more smooth, insensitive to outliers, and has a variable gradient with the difference value, instead of L1. In addition, in order to further improve image quality and detailed information, this paper also adds the structural similarity (SSIM) loss function, as shown in formula (2-4):

$$L_{cGAN}(G, D) = E_{x,y} [\log D(x, y)] + E_{x,z} [\log(1 - D(x, G(x, z)))]$$  \hspace{1cm} (2)

$$L_{SmoothL1}(G) = \begin{cases} 0.5 \times (y - G(x, z))^2, & \text{if } |y - G(x, z)| < 1 \\ |y - G(x, z)| - 0.5, & \text{otherwise} \end{cases}$$  \hspace{1cm} (3)

Where $E$ is the expected value, $x$ is the original image, $y$ is the target image, $z$ is the random noise, $G$ is the generator, and $D$ is the discriminator.

$$L_{SSIM(x,y)}(G) = 1 - \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{\mu_x^2 + \mu_y^2 + C_1(\sigma_x^2 + \sigma_y^2 + C_2)}$$  \hspace{1cm} (4)
\( \mu_x \) is the average brightness of the image in the x direction, \( \mu_y \) is the average brightness of the image in the y direction, \( \sigma_x \) is the standard deviation of the pixel value in the x direction of the image, \( \sigma_y \) is the standard deviation of the pixel value in the y direction of the image, and \( \sigma_{xy} \) is the covariance of the image pixel value, \( C_1 \) and \( C_2 \) are two constants, which can determine the specific dynamic range according to the data type of the image.

In summary, the generator objective function in this article is composed of CGAN, Smooth L1 and SSIM, as shown in the Equation 5:

\[
G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda_1 \mathcal{L}_{SmoothL1}(G) + \lambda_2 \mathcal{L}_{SSIM}(G)
\]  
(5)

In practice, our tests have revealed that the value setting of \( \lambda_1 = 100 \) and \( \lambda_2 = 5 \) works well. As shown in Fig.4, the improved loss function has the significant improvement in both convergence speed and accuracy.

![Fig. 4 The variation of loss function](image)

### 4 Experiments and analysis

#### 4.1 Data sources and experimental equipment

The experimental data in this paper is divided into two parts. First, the artificially synthesized sand-dust image data set is expanded by data enhancement methods such as random cropping, including 400 training images and 15 test images. Secondly, we download several natural sand-dust images without corresponding original images through the Internet as the data supplement for the test set.

The experimental computer is configured as the hardware device with the processor Intel Core i7, the main frequency of 3.60 GHz, the memory of 16 GB, and the graphics card Nvidia GeForce RTX 3070.

### 4.2 Experiment of artificially synthesized data set

In order to verify the effectiveness and advancement of the method proposed in this paper better, in addition to comparing with the original pix2pix\[^2\] model and two popular GANs (CycleGAN\[^22\] and GANILLA\[^23\]), we also introduced two state-of-the-art sand-dust image enhancement algorithms based on traditional methods (Shi et al.\[^9\]; Wang et al\[^5\]), and through the peak signal-to-noise ratio (PSNR\[^24\]) and structural similarity (SSIM\[^24\]) to quantify the improvement effect. PSNR is one of the most widely full-reference image quality assessment methods based on the difference between corresponding pixels. The larger the value of PSNR, the higher the image quality and the less distortion between the two images. SSIM is another important full-reference image quality assessment method, which evaluates the similarity between two images based on the perspective of brightness, contrast, and structure. The larger the value of SSIM, the higher the similarity between the two images.

It can be seen from Fig.5 that the images enhanced by the traditional enhancement method will have artifacts and poor saturation, while the sand-dust image enhancement algorithm based on deep learning performs very well in the color cast correction on the artificial synthetic data set. Among them, the pix2pix network’s color reproduction ability for sand-dust images is impressive, despite the blurring of foreground objects and loss of details. Our proposed algorithm further improves the color reproduction and at the same time enhances the texture detail information of the foreground target.

It is worth noting that the ablation experimental results are also presented in Table 1 and Fig.5. Although the pix2pix model far outperforms the traditional algorithm on the PSNR index of enhancing sand-dust images, the result is not good on the SSIM index. The addition of the CBAM module greatly improves the details of the image, especially the foreground objects, so the SSIM index of the enhanced image is also greatly improved accordingly. Nevertheless, the SSIM index after adding the CBAM module to the network is similar to the SSIM index of the optimal method in the traditional algorithm,
and compared with the pix2pix model, the PSNR index is not significantly improved. This is one of the reasons why we further improve the spatial attention and loss functions. The data in Table 1 show that our proposed method has the best performance on both the metrics of PSNR and SSIM, and is much higher than the current traditional sand-dust image enhancement algorithm.

### 4.3 Experiment of natural sand-dust image

In addition, to better verify the generalization of the proposed method, we also conducted enhancement experiments on natural sand-dust images.

It can be seen from Fig. 6 that in the natural sand-dust image enhancement experiment, although the method of Shi et al.\cite{9} has a relatively stable performance in enhancing detail information, the enhanced image has low saturation, color distortion, and blue artifacts in the background; the method of Wang et al.\cite{5} has high overall color reproduction, but sometimes red artifacts appear and details are blurred; CycleGAN\cite{22} and GANILLA\cite{23} do not work well; the color reproduction of the pix2pix model is acceptable, but the details are severely lost; in contrast, the sand-dust images enhanced by our algorithm have both high color reproduction and detailed texture information.
To further evaluate the proposed method, this paper also uses the non-reference image quality assessment method: the natural image quality evaluator (NIQE), which uses natural scene statistics in the spatial domain to evaluate degraded images. The lower values of NIQE represent the better image quality. As shown in Table 2, our method obtained the lowest NIQE values among all compared methods, which further demonstrates the superiority of our proposed method.

5 Conclusion

This paper proposes an improved pix2pix generative adversarial network for sand-dust image enhancement method, by introducing the dual attention mechanism into the generator of the UNet structure to enhance the local detail feature information of the main target in the foreground. Through the loss function of CGAN+SmoothL1+SSIM, the convergence accuracy and color reproduction are improved. In addition, we publish the first artificially synthesized sand-dust image data set online. The experimental results show that the enhancement method proposed in this paper has obvious advantages in both artificially synthesized images and natural real images.

In future work, we will improve the discriminator to increase image resolution and introduce unsupervised learning.
6 Declarations

- **Ethical Approval** Not applicable.
- **Competing interests** The authors declare that they have no conflict of interest.
- **Authors’ contributions** Zhongwei Hua wrote the main manuscript text. All authors reviewed and discussed the manuscript.
- **Funding** This research was sponsored by Shanghai Municipal Science and Technology Major Project No.2021SHZDZX0103 and Natural Science Foundation of Jiangxi Province No.20212BAB202026.
- **Availability of data and materials** All data generated or analysed during this study are included in this published article.

References


Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- githublink.zip