Research Developments in Generative Adversarial Networks for Image Restoration and Communication

Peng Zhou

Chengdu College of University of Electronic Science and Technology of China, Western High-Tech District, Chengdu 611731, China

Abstract. Generative Adversarial Networks (GAN) is always a popular study topic in artificial intelligence. This paper will analyze the principle of the GAN and introduce the development of the GAN and various derivative models. The improved Super-Resolution GAN (SRGAN) model and Cascading Residual Super-Resolution GAN (CR-SRGAN) model based on the GAN model achieve super-resolution of dark and old artifact images and solve the problem of color restoration and texture enrichment of dark and old artifacts. The GAN model is also widely used in the field of communication and information security. It proposes an End-to-End(E-to-E) communication encryption system based on Deep Convolutional GAN (DCGAN) to solve the secure transmission problem in wireless communication systems based on E-to-E learning. The system can realize encoding and decoding of input bits of arbitrary length with good generalization ability. Finally, the image restoration and communication encryption are summarized, along with an outlook on their development trends.

1 Introduction

Goodfellow proposed an emerging generative model (GAN) for unsupervised learning; by understanding and generating complex data distributions, GAN drives the development and application of deep learning techniques [1]. GANs, which originated from the concept of a two-player zero-sum game in game theory, is a training network similar to an adversarial game, learning by letting two neural networks play with each other and optimizing them iteratively until the model reaches a Nash equilibrium (NE) state [2]. GANs have been widely used in the fields of image restoration and communication encryption.

According to Zhu's research, GAN significantly improves image restoration. The original dataset was interpolated based on the dark old and noise processing to obtain the corresponding low-resolution images, and finally, through the training of the network model, to obtain the color restoration of super-resolution cultural relics images. This process dealt with the degradation of cultural relics images according to the characteristics of fading, dark old, and noise processing [3]. Domain information security, the research of An and others

Corresponding author: 1809040311@stu.hrbust.edu.cn

introduced CNNs into GAN to form DCGAN, and combined the generator of DCGAN as a cryptographic module with the CNN-based AE) to propose an encrypted e-to-e communication system due to DCGAN [4]. The system has a bit error rate (BER) capability similar to that of traditional digital modulation systems and can encrypt information in various forms, making it more difficult for illegal eavesdroppers to eavesdrop and decipher the signals.

This paper first introduces the basic structure of the GAN model, then introduces the GAN derivative model and its principle, and then discusses its potential applications in the fields of image restoration and communication encryption. Finally, the future development trends and research prospects of GAN are summarized.

2. Principles of the models

2.1 Principles of GAN

GAN consists of two neural networks: a generator(G) and a discriminator(D) [1]. The G's task is to generate realistic digital samples based on random noise, while the D is responsible for distinguishing real data from generated data. During the training process, the G tries to continuously enhance the quality of its generated samples to cheat the D, while the D corrects its mistakes by improving its recognition ability. This adversarial process creates a game where the G and the D promote mutually, enabling the G to generate increasingly realistic data. Eventually, after many iterations, ideally, the data generated by the G will be difficult to distinguish by the D, thus realizing high-quality data generation. The GAN structure is Displayed in Figure 1.

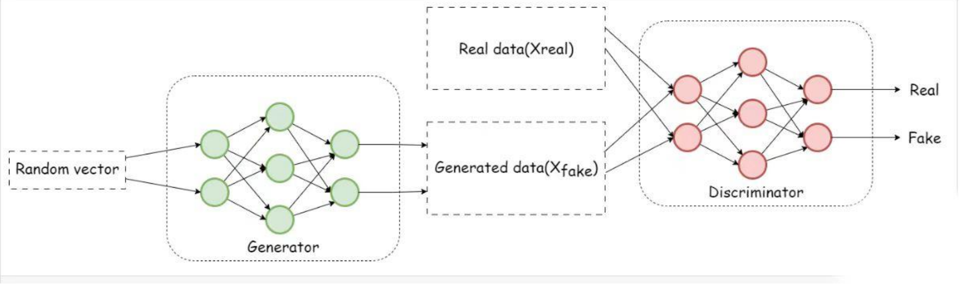

Fig. 1 GAN Architecture [5]

2.2 Principles of SRGAN

SRGAN is a deep learning model that uses GAN to recreate sharp, enlarged pictures. A D plus a G make up this system. Using convolutional layers, the G converts an input image with low resolution into an output image with high resolution. It is the D's responsibility to establish the authenticity of the generated high-resolution image. During training, D continuously improves its ability to create realistic images to cheat the D, resulting in higher visual fidelity. SRGAN also incorporates a perceptual loss function, which not only considers pixel-level differences but also emphasizes the similarity of the generated images in the feature space. This enhances the quality and details of the generated images. Overall, SRGAN's adversarial training mechanism enables it to generate more natural and clearer super-resolution images. The architecture of SRGAN is depicted in Figure 2.

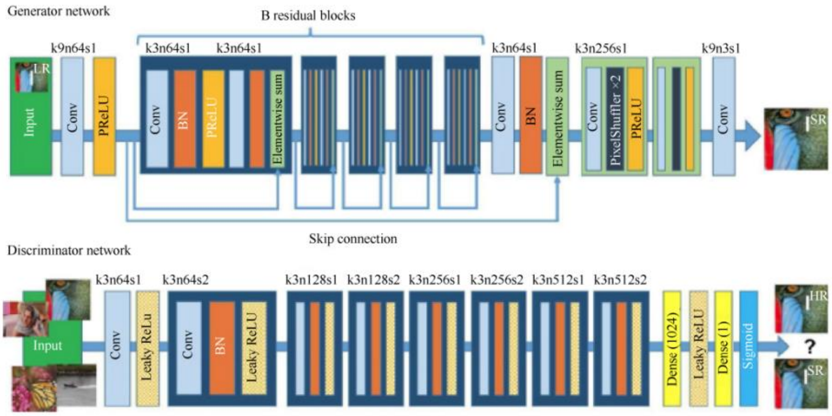

Fig. 2 SRGAN Architecture [6]

To enhance the realism of generated images, traditional single-frame image hyper-segmentation algorithms use Mean Squared Error (MSE) as the loss function. However, this approach tends to lose high-frequency texture details, resulting in less visually realistic generated images. To address this issue, SRGAN trains the generative model using a perceptual loss function. To improve the generated image's visual quality, this loss function highlights the perceptual differences between the generated and genuine images. The generative model's loss function has the following expression:

$$1^{SR} = l_X^{SR} + 10^{-3} l_{Gen}^{SR} \tag{1}$$

l_X^{SR} is the content loss and l_{Gen}^{SR} is the adversarial loss. The two loss functions add up to the perceptual loss [3].

2.3 Principles of DCGAN

DCGAN is a type of GAN specifically designed for generating images. DCGAN uses CNNs to create G and D. The G produces high-resolution images by gradually sampling random noise using an inverse convolutional layer [7]. Meanwhile, the D uses a convolutional layer to obtain features from input images and determine whether they are real or generated. During training, G and D engage in adversarial training, where the G aims to produce increasingly realistic images to fool the D, while the D's goal is to develop its ability to differentiate between real and generated images. The adversarial training mechanism allows DCGAN to generate high-quality images without the training instability and pattern collapse issues commonly found in traditional GANs.

The structure of DCGAN is illustrated in Figure 3.

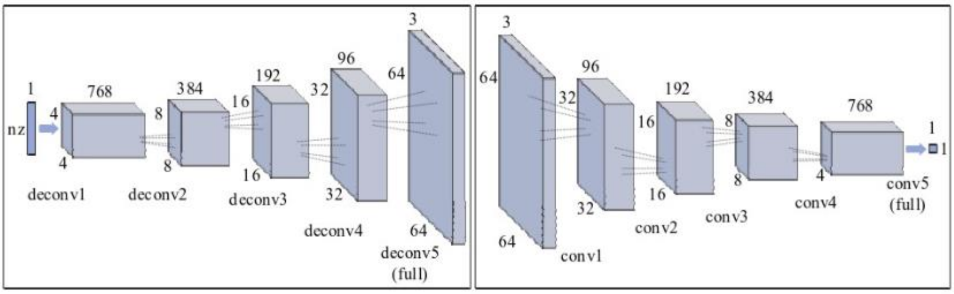

Fig. 3 DCGAN Architecture [8]

DCGAN addresses the issues of training instability, gradient vanishing, and model crashing that are prevalent in simple GAN to some extent through its unique structure. While DCGAN shares similar fundamentals with the simple GAN network model, it enhances the generative and discriminative networks by implementing them with a CNN network. This improvement aims to overcome the problems associated with simple GAN networks, ultimately enhancing the quality of the samples and expediting the network's convergence.

3 Application scenarios for GAN-derived models

3.1 Application of SRGAN in image restoration

Zhu has enhanced SRGAN and envisioned a CR-SRGAN for super-resolution(S-R) reconstruction of cultural relics and color restoration [3]. This differs from the traditional approach of training by interpolating and downsampling high-resolution images to obtain a low-resolution dataset. Instead, it processes the dark and fading characteristics of old cultural relic images, performs interpolation and noise processing on the original dataset to acquire the corresponding low-resolution image, and then trains the model to achieve super-resolution and color restoration of the relics.

The process of restoration of dark old cultural relic pictures using the CR-SRGAN is mainly separated into 4 steps:

1) Preprocessing the dataset: The dataset is preprocessed according to the characteristics of the cultural relic images to obtain corresponding high-resolution and low-resolution datasets.

2) Building the generative and discriminative networks: Only the high-resolution and low-resolution datasets acquired via standard downsampling are used to train the picture super-resolution model.

3) Training the CR-SRGAN using the generative and discriminative networks constructed in step 2, and using the datasets from step 1.

4) Reconstructing the high-resolution artifact image using the model generated from the training in step 3.

The architecture of CR-SRGAN is shown in Figure 4.

 https://doi.org/10.1051/itmconf/20257001020

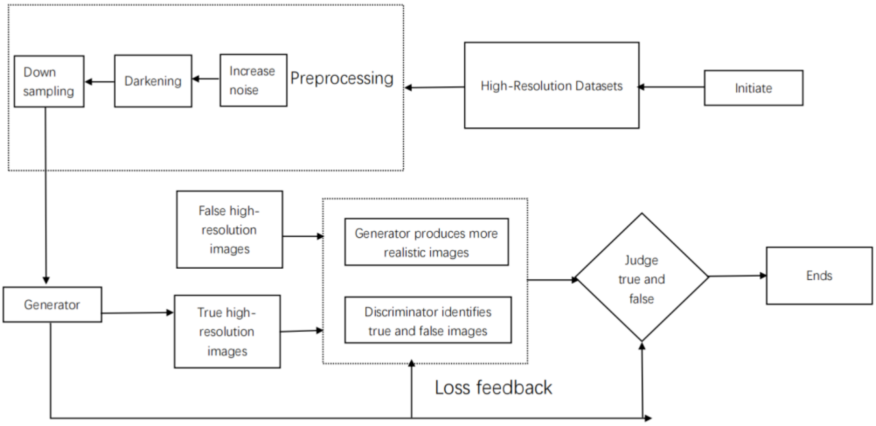

Fig. 4. CR-SRGAN Architecture [3]

Fig. 5 and Fig. 6 show the comparison between using “Bicubic”, “SRGAN”, “CR-SRGAN” and the original dark old image respectively.

(a) Dark old frescoes

(b) bicubic model

(c) SRGAN model

(d) CR-SRGAN model

Fig. 5. The comparisons of Bicubic; SRGAN; CR-SRGAN; Original dark old image [3]

(a) Dark old frescoes

(b) bicubic model

(c) SRGAN model

(d) CR-SRGAN model

Fig. 6. The comparisons of Bicubic; SRGAN; CR-SRGAN; Original dark old image [3]

3.2 Application of DCGAN in Communication

To tackle the secure transmission tests in e-to-e learning-based communication systems, researchers have introduced CNN into GAN to create DCGAN [4]. In this system, the G of DCGAN serves as the encryption module of AE based on CNN, and a secure, encrypted E-to-E communication system is proposed. Using CNN in this system offers the following advantages for E-to-E communication:

1) By employing DCGAN for encrypting information in long sequences, an encrypted E-to-E communication system is achieved, which is adaptable to different channels through E-to-E optimization. This system enables the encrypted transmission of signals in any long sequence.

2) The system is tested and examined on the Additive White Gaussian Noise channel and Rayleigh fading channel. The consequences indicate that the DCGAN-based encrypted E-to-E communication system outperforms traditional digital modulation systems.

3) By training an eavesdropping decoder to intercept the system E-to-E, it is confirmed that the system offers better eavesdropping security compared to the basic AE E-to-E communication system.

In traditional autoencoders (AE), the aim is to find a low-dimensional representation of the input figures in the hidden layer, allowing for minimal error when reconstructing the input at the output [4]. However, in AE-based E-to-E communication systems, the goal shifts to creating a signal representation that can withstand channel loss, ensuring better recovery of the original information at the receiver [9]. Unlike traditional AEs, which aim to remove redundant information to achieve a low-dimensional representation, AEs in E-to-E communication systems add redundant information to the input data to develop a more robust signal representation. The architecture of the AE is depicted in Figure 7.

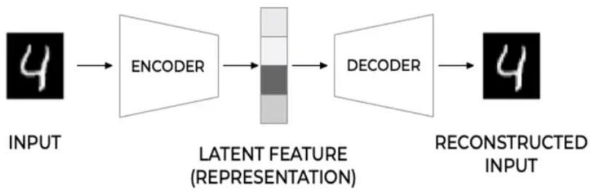

Fig. 7. Structure of AE [10]

4 Discussion

4.1 Discussion of better models for image-based restoration

Figures 5 (b), (c), and (d) display images of color reduction and (S-R) of dark ancient wall paintings using different ways, while (a) shows the original dark image. In Fig. 6, (b), (c), and (d) show images of color restoration and (S-R) of faded bronzes using different ways, with (a) showing the original dark old image.

The subjective evaluation method compares the reconstruction effect at the same magnification, primarily observing the texture information and color brightness of the generated images. It can be observed that the Bicubic model's enlarged super-resolution image is slightly blurred, while the SRGAN model produces clearer results, but the color remains dark and old. However, the method presented in this paper not only achieves clear texture but also enhances the color of the old cultural relics, thereby restoring the image to an older appearance. This is particularly evident in the crown of hair, eyebrows, and clothes of figures in Fig. 5(d), and at the top and middle of bronze ware in Fig. 6(d). The CR-SRGAN model reconstructs the image based on the original residual background color information, in line with how restorers would typically restore cultural relics by considering the residual background color. These results demonstrate that the CR-SRGAN model outperforms the comparative models.

4.2 Discussion based on a better model of E-to-E communication

Determining the quality of network learning is primarily dependent on the loss function. Inappropriate loss function selection will result in low model accuracy and other issues if the network structure stays the same. The loss function in this paper refers to equation (2). The GAN model is difficult to train, the D is too poor to provide an effective gradient, and too good will lead to gradient disappearance. Observing the loss functions of the GAN, CGAN, and DCGAN models for the Rayleigh channel under 16QAM modulation in Fig. 6, it can see that the GAN model produces a crossover point (i.e., reaches the NE) at 290 iterations, the CGAN model reaches the NE at 150 iterations, and the DCGAN model can reach the NE at 20 iterations.

The loss function is as follows:

$$\underset{G}{min}\underset{D}{max}V(D,G) = E_{i\sim p_r(i)}[\mathrm{lb}(D_D(i \mid c))] + E_{n\sim p_n(n)}[\mathrm{lb}(1 - D_D(G_G(n \mid c)))] \tag{2}$$

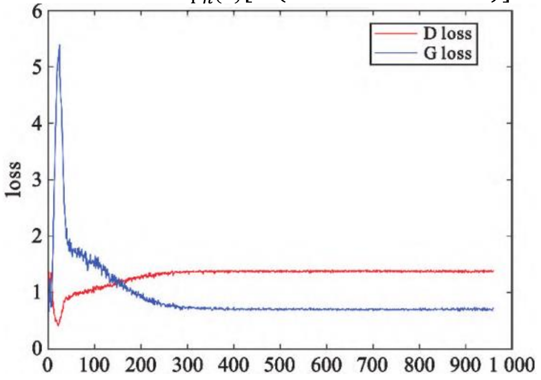

Fig. 8. CGAN-based loss function [11]

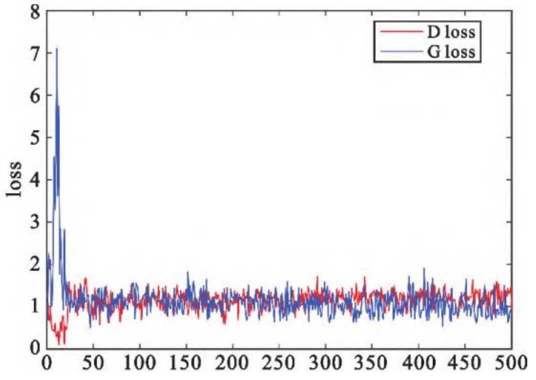

Fig. 9 DCGAN-based loss function [11]

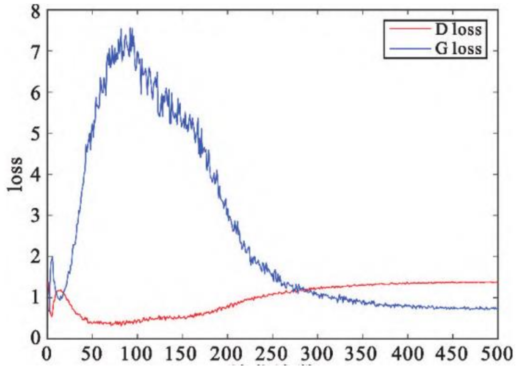

Fig. 10 GAN-based loss function [11]

After comparing Figures 8, 9, and 10, it is apparent that GAN has the slowest convergence, followed by CGAN. DCGAN demonstrates faster and more effective convergence compared to GAN and CGAN as the number of training instances increases. Analyzing different metrics, it can be concluded that the DCGAN method is most suitable for assessing channel modeling effects, allowing for intuitive and accurate judgment of the modeling effects of different methods. Furthermore, comparing the convergence effects of the loss function reveals that the DCGAN-based modeling method outperforms the CGAN and GAN-based methods.

5 Conclusion

This paper firstly introduces generative adversarial networks and their principles, and secondly introduces the principles and characteristics of two derivatives of the GAN model, the SRGAN model and the DCGAN model. Then the application of the SRGAN model and the improved CR-SRGAN model in image restoration is introduced, and an E-to-E communication system based on DCGAN design is also presented. Finally, it discusses and analyze the outstanding performance of the CR-SRGAN model in image restoration compared with other models and the advantages of DCGAN compared with other models in the field of communication. The GAN model plays an important role in many fields, and it is hoped that in the future, it can make greater breakthroughs in the perfect restoration of cracks, flaking, or limb mutilations of cultural relics, as well as to improve the safety of the E-to-E communication system. security of E-to-E communication systems.

References

1. G. I. J. Goodfellow, J. P. A. Pouget-Abadie, M. Mirza, et al., Generative adversarial nets, Proceedings of Annual Conference on Neural Information Processing Systems, New York: NIPS Press, pp. 2672-2680 (2014).
2. J. Yan, Research on strategy search algorithm in two-player zero-sum game problems, Dalian University of Technology, (2020).
3. Z. X. Lei, Q. Wu, Super-Resolution Reconstruction and Color Restoration of Cultural Relic Images Based on Generative Adversarial Networks, J. Xi'an Eng. Univ., 35(03), pp. 86-92 (2021).
4. Y. L. An, Z. R. Li, N. Li, et al., Design of encrypted end-to-end communication system based on DCGAN, J. Nanjing Univ. Posts Telecommun. (Nat. Sci. Ed.), pp. 1-9 (2024).
5. C. Ledig, L. Theis, F. Huszar, et al., Photo-realistic Single Image Super-resolution Using a GAN, (2016).
6. Y. Hu, Research on Mural Image Restoration and Super-resolution Reconstruction Method Based on GAN, North University of China (2023).
7. Z. Liu, Research on Transform Domain Image Restoration Technology Based on Deep Learning, Xi'an University of Electronic Science and Technology (2020).
8. X. A. Xu, G. Li, DCGAN-based image generation method for classroom expressions, Comput. Moderniz., (08), pp. 88-91+126 (2024).
9. C. Sun, C. Gao, Targeted genome duplications by amplification editing, Sci. China Life Sci., pp. 1-3 (2024).
10. U. Michelucci, An introduction to autoencoders, arXiv preprint arXiv:2201.03898, (2022).
11. F. F. Cheng, X. D. Wang, N. Wu, Design of end-to-end communication system under DCGAN channel, Telecommun. Technol., 62(06), pp. 742-748 (2022).