DOI: 10.54254/2755-2721/106/20241263

Imagine Denoising Methods Based on GANs

Zhenqi Li

Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, 200030, China

usjjei13@sjtu.edu.cn

Abstract. With the continuous development of technology, people's demand for image clarity is also constantly increasing. Whether in the fields of medical imaging, aerospace, or people's daily lives, noise in images seriously affects their clarity. Therefore, how to remove noise on the premise of preserving image details has now become a hot research topic. In the domain of image denoising, both early spatial domain filtering methods and recently proposed convolutional neural network models have certain limitations. Compared to other denoising methods, GANs can better remove noise from images and improve image quality. Therefore, this article summarizes and organizes image denoising methods based on GAN models. This article explains four GAN based image denoising methods, namely GAN, WGAN, DNGAN, and GCBD, from the perspectives of framework structure, advantages and disadvantages, and application fields. At the same time, this article also analyzes the development trend of GAN application in image denoising and makes prospects.

Keywords: GAN, Image denoising, Deep learning.

1. Introduction

With the development and progress of technology, people's demand for clearer and more accurate images is also increasing day by day. However, images captured through electronic devices are inevitably created with noise. Noise in images often manifests as isolated pixels or pixel blocks which have significant visual impact on people, disrupting the actual information content of the image and making it unclear. Therefore, how to reduce noise without destroying the original features of the image has become a key research topic at present.

Early image denoising methods were mainly achieved through simple spatial domain filtering, such as mean filtering, median filtering, etc. Subsequently, with the emergence of wavelet theory, image denoising methods in the transform domain showed some advantages and improved in preserving some edge details in the image compared to spatial domain methods. In recent years, with the improvement of the performance of graphics processing hardware devices, deep learning and other approaches have yielded significant success in the field of image denoising. For example, Convolutional Neural Network models can solve denoising problems through independent learning. However, for this method, the design of the network structure and the setting of various hyperparameters are still empirical. Meanwhile, there is no pair of real noise and noise-free data [1]. Therefore, there is still room for improvement in this method.

DOI: 10.54254/2755-2721/106/20241263

In recent years, the computing power of graphics processing devices has continuously improved. The performance of Convolutional Neural Network models generated by two-dimensional data operations has also been greatly improved. As a result, the CNN models are widely used in the field of image processing. Convolutional models such as LeNet form different network models by using convolution modules of different sizes, quantities, and arrangements, and can complete a large number of digital image processing tasks through training and learning. Traditional filtering denoising methods use ensemble filters to compute images for denoising, while in CNNs, there is also a convolutional structure with similar functionality. Therefore, methods based on convolutional neural networks can also achieve denoising.

Compared with traditional methods, the setting of filters in CNN relies entirely on automatic training and learning of the model, rather than manual design. Therefore, although traditional filtering methods are still effective in some simple scenarios, CNN has shown superior performance in image denoising due to its strong adaptability, better robustness, and higher processing efficiency.

However, traditional deep learning based denoising models, such as the convolutional neural network mentioned earlier, require training with noisy images as training samples and the corresponding image's noiseless state as sample labels. These models also need to subtract the output image of the neural network model from its corresponding label as a loss function for backpropagation to correct the network parameters. The performance of such models is highly dependent on the diversity and quality of the training set, so they may perform poorly when faced with unseen types or structures of noise. Meanwhile, most of them are designed for specific tasks and lack flexibility. When dealing with various types of noise and image features, it may be necessary to redesign the network structure. Therefore, researchers have begun to attempt to apply Generative Adversarial Networks (GANs) to the field of image denoising. During adversarial training, the generator learns many complex distributions, which enables GANs to better adapt to different noise patterns and image features. At the same time, GANs can be trained unsupervised, which makes it very useful in scenarios where data annotation is scarce.

In this paper, session 1 introduces the research background and current status of image denoising methods based on GANs. Session 2 introduces the denoising principles and model structures of four image denoising methods, including GAN, WGAN, DNGAN, and GCBD. Session 3 provides a brief evaluation of the advantages, disadvantages, and application scenarios of the four image denoising methods mentioned in session 2. It also discussed the prospects and possible improvement directions of image denoising methods based on GANs. Session 4 summarizes the entire paper and provides prospects for the future of using GANs for image denoising. This article reviews the research progress of image denoising based on GANs and puts emphasis on analyzing the four methods mentioned above, aiming to provide reference and inspiration for future researchers.

2. Image denosing methods based on GANs

Classic image denoising methods, including spatial domain filtering, variational denoising, total variation regularization, non local regularization, sparse representation, and so on, all have their own problems such as difficulty in parameter adjustment, edge blurring, and loss of texture details [2]. Meanwhile, most of these methods are supervised methods, and in real-world scenarios, the continuous challenge of obtaining clean noise pairs for supervised methods remains daunting [3]. Therefore, researchers have proposed various image denoising methods based on GANs and achieved excellent results.

2.1. GAN denoising principle

Nowadays, GAN has had a wide range of applications in medical image denosing, low light image denoising, video denoising and other fields. In the process of using GAN for image denoising, the generator learns to minimize the divergence between the generated image and the genuine image. Its duty is to convert the noise into a clear image, and then generates the output as true and clean as possible from the noisy or contaminated input. The discriminator, on the other hand, continuously learns to distinguish whether the input image is a clear image from a real dataset or a denoised image. It is responsible for evaluating the quality of the generator. The structure of GAN is shown in the Figure 1.

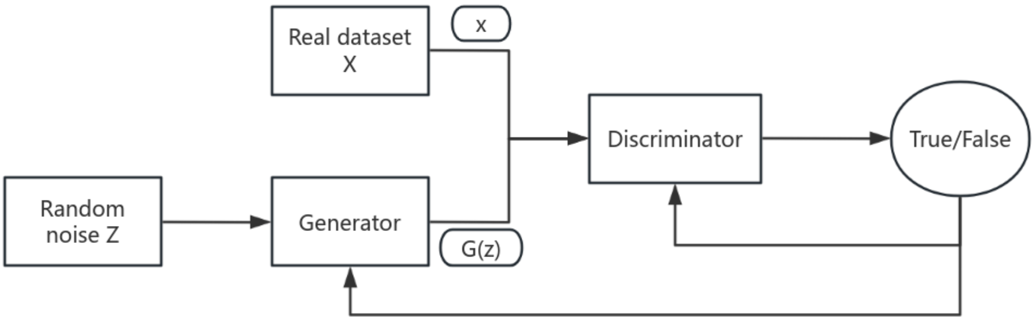

Figure 1. The structure of GAN.

The loss function of the generator can be based on the output of the discriminator, while incorporating some indicators related to image quality (such as perceptual loss, total variational loss, etc.). The objective function of GAN is as the formula (1).

$$\min_{G}\max_{D} V(D,G) = E_{x\sim pdata(x)}[log\, D\,(x)] + E_{z\sim p_z(z)}[log(\,1 - D(G(z)))] \tag{1}$$

where G represents the Generator, D represents the discriminator. $pdata(x)$ is the distribution of real data, and $p_z(z)$ is the distribution of random noise. $E_{x\sim pdata(x)}[log\, D\,(x)]$ and $E_{z\sim p_z(z)}[log(\,1 - D(G(z)))]$ respectively represents the expected loss of the discriminator on real data and generated data [4].

2.2. WGAN denoising principle

Though GAN can generate clear images in image denoising process, it still has some issues, such as instability and pattern collapse during training. To address the issues, Wasserstein GAN (WGAN) was created. WGAN uses Wasserstein distance (also known as Earth movement distance) to assess the distinctions between the generated distribution and the true data distribution. Compared to traditional JS divergence and KL divergence, Wasserstein distance has better mathematical properties and stability, which can make the training process smoother. In WGAN, the goal of the generator is to minimize the Wasserstein distance, which means that the distance between the generated image and the real image should be as small as possible. What is more, unlike the GAN, the discriminator of WGAN is not simply judging the authenticity of the image, but gives a real value representing the distance between the input and the real distribution. This design makes the model more expressive. In order to ensure that the discriminator of WGAN satisfies the local smoothness of the function, WGAN introduces the method of weight clipping. In this way, WGAN can ensure that the output of the discriminator is not extreme, thereby avoiding instability during the training process. The loss function of WGAN can be formulated as:

$$\min_{G}\max_{D} V(D,G) = E_{x\sim p_r} D(x) - E_{y\sim p_n} D(G(y)) - \lambda(||\nabla_z D(z)||_2 - 1)^2 \tag{2}$$

where $E_{x\sim p_r} D(x) - E_{y\sim p_n} D(G(y))$ indicate the Wasserstein distance estimation. Meanwhile, to regularize the discriminative subnetwork D, $\lambda(||\nabla_z D(z)||_2 - 1)^2$ is set to work as a gradient penalty term.

Compared to GAN, WGAN is more stable and can better capture data distribution. Meanwhile, WGAN can generate higher quality images in denoising tasks, and reduce artifacts and distortion. Research has indicated that training frameworks utilizing WGAN are capable of effectively addressing the issue of image blurring. They can also achieve superior denoising outcomes compared to existing denoising methodologies.

2.3. DNGAN denoising principle

The structure of DNGAN is similar to GAN, but it focuses more on denoising. DNGAN typically combines multiple loss functions such as adversarial loss and reconstruction loss (such as using L1 or L2 norm to evaluate the difference between the generated image and the target clear image, thereby improving denoising performance). Compared with GAN, DNGAN can more effectively capture and model different types of noise through a specially designed network architecture, better understand the characteristics of noise, and achieve higher quality denoising. In DNGAN, the generator is trained to map the noisy image to the ground truth. In the meanwhile, the discriminator works as the loss function. Its duty is to assess the discrepancies between the generated output and the ground truth. Research shows that compared with other popular methods in image denoising, the DNGAN method can achieve a better denoising effectiveness [5]. DNGAN also has a wider range of application scenarios. It can be used for processing medical images and mobile phone photos. It can even be used for denoising video frames, significantly improving the viewing experience of videos.

2.4. GAN-CNN Based Denoising (GCBD)

In many cases, the specific information of noise is unknown. In these cases, blind denoising methods are used. However, classic image blind denoising methods such as BM3D mostly rely on human prior knowledge, which makes it difficult for them to extract the characteristics of the image completely. Meanwhile, most methods can merely utilize the internal information of the input image without using any external information. And the existing models used in CNNs based discriminative learning methods are designed to remove known noise and cannot achieve good results in blind denoising [6]. Therefore, Jingwen Chen et al. proposed a new training framework, the GCBD method: first, train a GAN to learn the noise distribution present in the input noisy image and generate noise samples. Next, the noise blocks gathered during the first step are utilized to create the corresponding training dataset., and finally, a deep CNN is trained to denoise the given noisy image. This approach enables generative networks to learn the mapping from clean images to images that possess noise characteristics similar to the provided data. Therefore, it can address some of the shortcomings of methods such as BM3D. The framework of GCBD is shown as Figure 2.

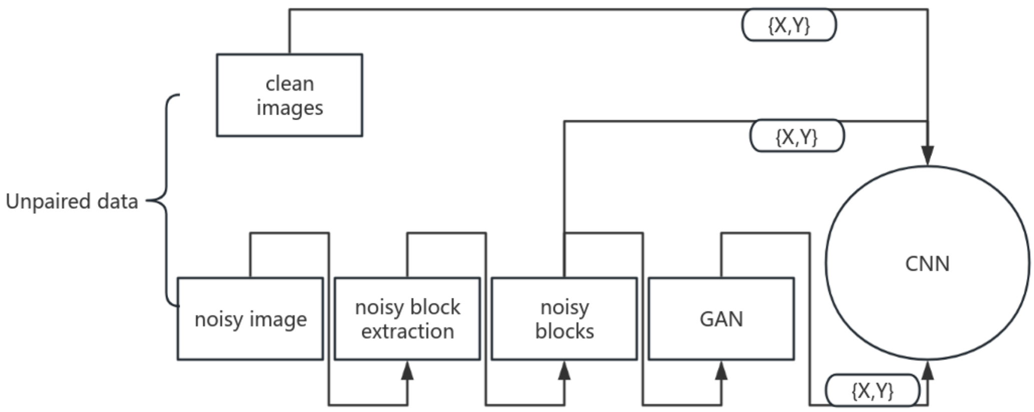

Figure 2. The structure of GCBD.

where unpaired data is given, and approximate noise blocks from noisy images are exploited to train a GAN. After training the model, some noise blocks are sampled from it. To obtain paired training data, these sampled noise blocks and the extracted noise blocks are then combined with clean images. After this, the obtained data is used for training a CNN, which works to denoise the input images.

DOI: 10.54254/2755-2721/106/20241263

3. Evaluation of various GAN-Based image denoising methods

In the previous section, this paper provided a detailed introduction to the application of models such as GAN, WGAN, DNGAN, and GCBD in image denoising. This section will comprehensively evaluate the advantages and disadvantages of these methods. It aims to provide readers with a comprehensive understanding and reference for selecting appropriate methods for subsequent research and application.

3.1. Evaluations of GAN on image denoising

GAN are usually able to generate clearer and more natural images in denoising tasks. Through adversarial learning, the generator can learn better data distribution, which enables it to preserve the details of the image more efficiently. It can also be trained for removing different types of noise like Gaussian noise, to achieve good generalization ability. This makes GAN perform well in various image denoising tasks.

However, even though GAN have many advantages in image denoising, it will come into problems when dealing with large dataset. That's because its training process is often complex and computationally intensive, requiring a large amount of data and time, and may encounter the problem of model collapse. When exhibiting signs of model collapse, even if the GAN are capable of generating images across the entire domain, the limited variety of the images would still hinder denoising efforts [7].

3.2. Evaluations of WGAN on image denoising

WGAN can tackle the blurriness problem linked to pixel-wise loss functions and boosts GAN performance. In this way, it produces denoised images with clearer and more realistic texture details. [8]. Nevertheless, it is still sensitive to weight clipping and hyperparameter selection. In addition, despite the introduction of Wasserstein distance, WGAN still requires relatively high computational resources for training. To address these issues with WGAN, researchers have combined WGAN with the Pix2Pix model to develop a new method that combines effectiveness and generalization in the field of image denoising [9].

3.3. Evaluations of DNGAN on image denoising

Compared with GAN, DNGAN's structure and loss function are optimized for image denoising, making it perform well in handling noise and thus more accurately removing noise. And DNGAN can handle different types of noise (such as Gaussian noise, salt and pepper noise, etc.), demonstrating strong denoising capabilities under various noise conditions. Meanwhile, DNGAN pays special attention to preserving the details and structural information of the image during denoising, which makes it outstanding in processing images with complex textures.

Despite DNGAN's outstanding performance, further optimization is needed in terms of training complexity and computational resource consumption to better apply it to practical scenarios.

3.4. Evaluations of GCBD on image denoising

In blind denoising problems, GCBD has shown considerable performance in denoising Gaussian noise, mixed noise, and real-world noise. Due to the unavailability of paired training datasets, methods such as DnCNN-B cannot work well. However, to solve the problem of the lack of training data, GCBD brings in GANs. In this way, GCBD can effectively estimate the noise distribution and expand the training data. Finally, it can achieve more significant denoising effects. In addition, GCBD does a pretty good job at retaining details, such as the spark of the light [6]. But this method still has some limitations, for example, it assumes the noise to be additive. Moreover, noise is also assumed to have a mean of zero. The performance of GCBD under other noise conditions needs further research.

DOI: 10.54254/2755-2721/106/20241263

3.5. Overall evaluations

In order to clearly demonstrate the advantages and disadvantages of each method in image denoising, Table 1 summarizes the performance and main characteristics of GAN, WGAN, DNGAN, and GCBD in different application scenarios.

As shown in the table, all four methods mentioned above have their own shortcomings. However, it cannot be denied that they have excellent performance in image denoising in certain scenarios.

Table 1. A demonstration of different image denoising methods.

Methods	Advantages	Disadvantages
GAN	• Preserve details of image • Address different types of noise • Generate cleaner images	• Complex and intensive train process
WGAN	• Address the blurriness problem • Generate cleaner images	• Sensitive to weight clipping and hyperparameter selection
DCGAN	• Remove noise more accurately • Address different types of noise • Handle images with complex textures	• Complex and intensive train process
GCBD	• Solve the problem of lack of training data • Perform well at retaining details	• The performance under other noise conditions needs further research

GANs based image denoising methods have made significant progress in the fields of computer vision and image processing in recent years. GANs have become an important technique in the field of image denoising due to its powerful generation ability and the ability to learn data distribution. It has been applied in removing noise from Monte Carlo rendered images, tomography images, and real images [10]. In addition, image denoising techniques based on GANs, such as WGAN, DNGAN, GCBD, etc., have also demonstrated good denoising effects. Although these technologies still have limitations such as high computational resource requirements, they provide excellent references and potential research directions for the field of image denoising. Future research can start from the following directions: how to construct a comprehensive GAN model to cope with noise in different situations; How to reduce the complexity of the training process.

4. Conclusion

In conclusion, this study introduces four image denoising methods based on GAN, WGAN, DNGAN, and GCBD. Explained their principles and model structures. Meanwhile, this article also discusses the characteristics of each of these algorithms and summarizes their advantages and disadvantages. The article also discusses the application scenarios of these methods.

In the future, researchers can explore the combination of GANs and other deep learning techniques such as variational autoencoders, graph neural networks, etc. to achieve stronger denoising capabilities. As mentioned in this article, the GCBD method combines GANs with CNN for image denoising and has achieved significant results in the field of blind denoising. In addition, we can also focus on improving the network architecture of GANs to better utilize the dataset and reduce the pressure on computing resources when used in the field of image denoising.

With the improvement of computing power and further optimization of algorithms, GANs based denoising technology is expected to achieve clearer and more efficient image processing. By then, GAN based image denoising technology is expected to be widely applied in various related fields such as medical image denoising, autonomous driving, video processing, etc. Research in this area can have a certain promoting effect on the progress of related technologies and people's lives.

DOI: 10.54254/2755-2721/106/20241263

References

[1] Wang X, Wu K, Zhang Y, Xiao Y and Xu P 2023 A gan-based denoising method for chinese stele and rubbing calligraphic image. *The visual computer*.

[2] Fan L, Zhang F, Fan H, and Zhang C 2019 Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine and Art*.

[3] Niu C, Li K, Wang D, Zhu W, Xu H and Dong J 2024 Gr-gan: a unified adversarial framework for single image glare removal and denoising. *PATTERN RECOGN*.

[4] Hu X and Zhao W 2024 Multi-scale information fusion generative adversarial network for real-world noisy image denoising. *MACH VISION APPL*.

[5] Chen Z, Zeng Z, Shen H, Zheng X, Dai P and Ouyang P 2020 Dn-gan: denoising generative adversarial networks for speckle noise reduction in optical coherence tomography images. *BIOMED SIGNAL PROCES*, 101632.1-10.

[6] Chen J, Chen J, Chao H and Yang M 2018 Image blind denoising with generative adversarial network based noise modeling. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE.*

[7] Chen C 2020 Generative adversarial networks for electron miscroscope image denoising.

[8] Chen S, Shi D, Sadiq M, and Cheng X 2020 Image denoising with generative adversarial networks and its application to cell image enhancement. *IEEE ACCESS*, 82819-31.

[9] Tirel L, Ali A M and Hashim H A 2024 Novel hybrid integrated pix2pix and wgan model with gradient penalty for binary images denoising.

[10] Paliwal A, Zeng and Kalantari N K 2021 Mask dngan: multi-stage raw video denoising with adversarial loss and gradient mask.