

A comparative Analysis of Image Enhancement Using Autoencoders and Generative Adversarial Networks

Minhua Ying*

School of Southeast University Cheng Xian College, Nanjing, China

*Corresponding author: 219121201@cxxy.seu.edu.cn

Abstract. The purpose of this study, which focuses on image enhancement methods, is to investigate how various methods contribute to higher-quality photos in particular applications. Tracing the history of image improvement, encompassing the technological evolution of early photography from the 19th century to the present. The foundation of conventional image enhancement technology, which directly affects image elements and transformation coefficients, respectively, emphasizes frequency domain and spatial domain processing. It is noted that the application of autoencoders and generative adversarial networks to image enhancement as a deep learning training strategy has shown effectiveness in picture classification, target recognition, and image segmentation. This article primarily examines and contrasts autoencoders and generative adversarial networks, summarizing their applicability, constraints, benefits, and drawbacks and talking about the patterns in their upcoming development. The study has guiding value for advancing the development of picture technology and offers significant theoretical and empirical support for the field of image enhancement.

Keywords: Image enhancement; Autocoders; GAN.

1. Introduction

Improving the quality of images for particular applications is the goal of image enhancement, one of the most significant technologies in the field of image processing. Generally speaking, the fundamental idea behind image enhancement is to change a picture's information contribution to make it more appropriate for a given use case, as well as to provide the model with greater learning invariance to improve its capacity for generalization and lower the likelihood of overfitting [1, 2].

The history of image enhancement is closely tied to the development of imaging technologies and the need to improve the quality of visual information for various applications. It mainly went through the following stages: Early Photography in the 19th Century, Film Photography and Darkroom Techniques in the 20th Century, Introduction of Digital Imaging in the 1960s-1970s, Computer Vision and Image Processing in the 1980s-1990s, Advancements in Medical Imaging in the Late 20th Century, Digital Photography and Consumer Applications in the 2000s-Present.

Image enhancement is an effective training strategy that can increase the image collection of neural network models without requiring additional images, and has achieved good results in deep learning such as image classification, target detection, and image segmentation [2]. Various techniques play a major part in improving images. The primary foundation of conventional picture enhancement techniques is frequency domain and spatial domain processing. Image improvement in the spatial-based realm works directly with picture elements. The primary benefit of spatially-based domain approaches is their decreased complexity and ease of implementation. However, the phrase "frequency-based domain image enhancement" refers to the examination of mathematical signals or functions that have a frequency component and that directly alter the image's transform coefficients, such as the Fourier transform, Discrete wavelet transform (DWT), and Discrete cosine transform (DCT). The transform coefficients are manipulated for this technique to function. Low computational complexity, ease of seeing and modifying the image's frequency composition, and the simplicity with which unique transformed domain attributes can be applied are the benefits of frequency-based image enhancement [3]. Image enhancement has been applied in many fields, including computer vision, engineering, mathematics, and astrophysics [2].

In this paper, Autoencoders and Generative Adversarial Networks methods are summarized, and their applicability, limitations, advantages, and disadvantages are analyzed to further summarize the contributions and existing problems of the two methods to image enhancement and analyze the future development trend of the two methods. First, the paper focuses on the characteristics of the two methods. In the second part of this chapter, the comparative analysis of Autoencoders and Generative Adversarial Networks methods is mainly discussed. The Autoencoders and Generative Adversarial Networks methods are discussed in the third part of this chapter.

2. Methods

2.1. Autoencoders

As seen in Figure 1, an autoencoder is a neural network made up of an encoder and a decoder, each with its own set of learnable parameters. The encoder's job is to map an input (x) into a latent encoding space (which can hold any kind of data, including text, audio, video, and 2D or 3D pictures) to produce a latent code h [4]. An MRI slice of 256×256 voxels, for instance, can be converted by the encoder into a latent code vector with a size of 50×1 (or any other desired size). Conversely, the input data is extracted from the code by the decoder. In the same example, the decoder attempts to recreate the original 256×256 MRI slice, x' , using the encoder's output, h (vector size 50×1) [4].

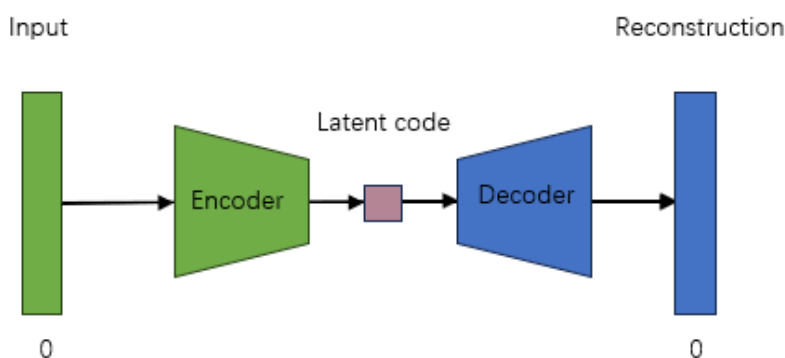


Fig. 1 The structure of the autoencoder [4].

The encoder and decoder neural networks make up the autoencoder's structure. After the encoder processes the input, a latent code is created. The decoder uses the latent code to produce the output. The architecture of the decoder in most autoencoders is a replica of the encoder's architecture. The network wants to provide an output that is the same as the input. In this instance, the autoencoder receives an image of the number "0" and uses the latent coding it has learned to recreate the image.

2.2. Generative Adversarial Networks (GAN)

A discriminant network and a generation network make up the GAN's network structure. In Figure 2, the model structure is displayed [5]. After receiving the random variable z , generator G creates a fictitious sample of data $G(z)$. The generator's goal is to produce samples that are as near to the originals as feasible. The actual data (x) and the data produced by the generator $G(x)$ make up the discriminator D 's input. Typically, it produces a probability value as an output, representing the likelihood that D will conclude that the input is a real distribution. Output 1 if the input is true data; else, output 0. Simultaneously, G will receive the discriminator's output back to direct G 's training. In an ideal world, D is unable to distinguish between produced data $G(z)$ and real data x ; in other words, D 's output probability value is always $1/2$ (equivalent to random guessing), which is the same as random guesswork, and the model reaches its optimal value at this point. In practical applications, deep neural networks are usually utilized to generate generative and discriminative networks.[5].

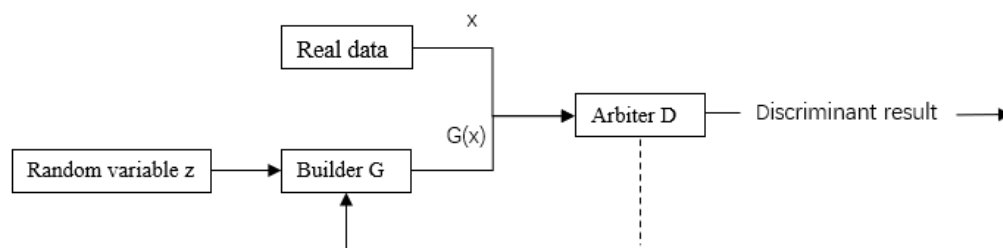


Fig. 2 GAN network model structure diagram [5].

The foundation of generative adversarial networks is a game between two machine learning models, usually neural networks, in the sense of game theory [6].

3. Comparative Analysis

In the field of image enhancement, both autoencoders and generative adversarial networks are widely used, but there are some differences in their methods and effects. This section compares and examines the two approaches in objectives and tasks, structure, and Assessment of image quality.

3.1. Objectives and Tasks

Learning a compact representation of the input data is the primary objective of an autoencoder. It is made up of an encoder and a decoder that work together to minimize the reconstruction error between the input and the decoded output to learn a low-dimensional representation of the data.

Generative adversarial networks' main goal is to train a network of generators to produce samples that closely resemble real data. The generator and discriminator in a GAN work together to generate increasingly realistic samples as the player progresses through the game, while the discriminator attempts to separate the generated samples from the real ones.

3.2. Structure

The structure of Autoencoders consists of two main parts: encoders and decoders. The encoder maps input data into a low-dimensional space of possible representations after receiving the data, which is often a vector or matrix. This mapping can be done using neural network layers, such as convolutional and fully connected layers. Typically, a latent representation or encoding is the result. The encoded latent representation is received by the decoder, which then maps it back to the original input space.

The structure of a generative adversarial network consists of two main parts: generator and discriminator. The generator is in charge of creating visuals that mimic actual data from random noise or latent space. A neural network, which might be a fully connected layer, a convolutional neural network (CNN), or another structure, often makes up the generator's structure. The discriminator is in charge of differentiating between the generator's produced images and actual data. Its objective is to make the images produced by the generator increasingly lifelike so that the discriminator finds it challenging to tell them apart.

3.3. Assessment of Image Quality

To evaluate the quality of an image, autoencoders often compare the differences between the input and reconstructed images. This is calculating the difference between the original image and the reconstructed image using a loss function. Since the objective of Generative adversarial networks is to produce realistic images, evaluating the image quality of these images quantitatively is difficult due to the subjectivity of image quality. So Generative adversarial networks lack an objective function, which makes it difficult to compare the performance of different models [7].

So Autoencoders are generally more straightforward in terms of directly reconstructing input images, while Generative adversarial networks are better at generating high-quality, diverse, and realistic images.

4. Discussion and Prospects

4.1. Applicability

The primary uses of automatic encoding in picture enhancement include feature learning, data augmentation, and noise reduction and deblurring. Super-resolution, picture translation, style conversion, and data production are the primary uses of GANs in image enhancement.

Selecting the best technique for picture enhancement depends on the nature of the data and the task at hand. Thus, autoencoders are better suited for jobs that prioritize picture reconstruction, while GAN performance is better suited for tasks that produce a wide range of high-quality images.

4.2. Limitations

Autoencoders tend to generate an average representation of the input training data and lack generative diversity. In a supervised learning setting, autoencoders require large amounts of labeled data, while in unsupervised or semi-supervised cases, good latent representations may not be learned. Unrealistic sample generation and mode collapse are issues that generative networks face [8]. The quality of images generated by Generative adversarial networks is difficult to quantitatively assess because the goals of the generator and discriminator do not directly correspond to the quality of the generated images.

4.3. Future Works

As two methods of image enhancement, autoencoders, and generative adversarial networks have been continuously improved. Now they are the most promising generative models, and there are more detailed plans for future work.

The model was first created to reduce dimensionality, but it has since enabled the creation of the first deep learning models, revolutionizing the area of representational learning. Along with generative adversarial networks, it is currently one of the most promising generative models. Autoencoders are anticipated to become increasingly popular in the future years in the field of brain disease research for two main reasons. First of all, these methods appear capable of overcoming the limitations imposed by feature engineering. Second, since human and animal learning is primarily unsupervised—that is, learning the structure of the world by observation rather than being given the category of each object—the potential utility of strictly supervised learning is probably limited. Consequently, it is anticipated that additional study will concentrate on the kind of learning—unsupervised learning—that the autoencoder makes possible. This kind of learning still has a lot of promise because it is still largely unexplored in general and in studies on brain illnesses in particular [9]. The inadequacy of the GAN essential hypothesis prevents GAN models from producing superior generative models. Therefore, the most important application for future work is to make breakthroughs in theoretical areas to address problems, for example, difficulties in training, non-union, and model breakdown [10]. Future advancements are still required even with certain generally better tactics, such as weights trimming and regularization, Nash equilibrium, and new loss functions [11-15].

5. Conclusion

Image enhancement plays an important role in image processing. In this paper, first, an overview of image enhancement is given, and two image enhancement methods are introduced, namely Autoencoders and Generative Adversarial Networks. Then through the comparative analysis and discussion of the Autoencoders and Generative Adversarial Network methods, the role and

development of the a and g methods are highlighted. Autoencoders and generative adversarial networks are two important generative models in deep learning, and they have shown a wide range of applications in many fields.

Autoencoders are widely used for feature learning and data reconstruction in image processing. In terms of anomaly detection, by learning representations of normal data, autoencoders can detect data points that differ from normal patterns. In natural language processing, autoencoders are also used in word embedding and text generation.

Generative adversarial networks have excellent performance in image generation, image enhancement, and style transfer. In the field of art generation, generative adversarial networks can create realistic works of art, demonstrating their great potential in the creative field.

Anyway, in the future, research on the interpretability, robustness, and multi-modal learning of these two models will promote their deeper applications in medical care, art creation, intelligent interaction, and other fields.

References

- [1] Qi Y, Yang Z, Sun W, et al. A comprehensive overview of image enhancement techniques. *Archives of Computational Methods in Engineering*, 2021: 1-25.
- [2] Shukui S, Jing F, Zhongqiang S, et al. A Review of image data enhancement based on Deep Learning. *Computer Science*, 2024, 51(01): 150-167.
- [3] Singh G, Mittal A. Various image enhancement techniques-a critical review. *International Journal of Innovation and Scientific Research*, 2014, 10(2): 267-274.
- [4] Pinaya W H L, Vieira S, Garcia-Dias R, et al. Autoencoders. *Machine learning*. Academic Press, 2020: 193-208.
- [5] Liang Junjie, Wei Jianjing, Jiang Zhengfeng. A review of generative adversarial networks. *Computer Science and Exploration*, 2020, 14(1): 17.
- [6] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Advances in neural information processing systems*, 2014, 27
- [7] Denton E L, Chintala S, Fergus R. Deep generative image models using a laplacian pyramid of adversarial networks. *Advances in neural information processing systems*, 2015, 28.
- [8] Liu Z S, Siu W C, Chan Y L. Photo-realistic image super-resolution via variational autoencoders. *IEEE Transactions on Circuits and Systems for video Technology*, 2020, 31(4): 1351-1365
- [9] Saxena D, Cao J. Generative adversarial networks (GANs) challenges, solutions, and future directions. *ACM Computing Surveys (CSUR)*, 2021, 54(3): 1-42.
- [10] Salimans T, Goodfellow I, Zaremba W, et al. Improved techniques for training gans. *Advances in neural information processing systems*, 2016, 29.
- [11] Aggarwal A, Mittal M, Battineni G. Generative adversarial network: An overview of theory and applications. *International Journal of Information Management Data Insights*, 2021, 1(1): 100004.
- [12] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. *International conference on machine learning*. PMLR, 2017: 214-223.
- [13] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of wasserstein gans. *Advances in neural information processing systems*, 2017, 30.
- [14] Kodali N, Abernethy J, Hays J, et al. On convergence and stability of gans. *ar**v preprint ar**v:1705.07215*, 2017.
- [15] Mao X, Li Q, **e H, et al. Least squares generative adversarial networks. *Proceedings of the IEEE international conference on computer vision*. 2017: 2794-2802.