Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 7, No. 2, 2023

Lung Nodule Data Enhancement Algorithm Based on Generative Adversarial Network

Yunjiao Ma

School of Information, Yunnan University of Finance and Economics, Kunming Yunnan, 650221, China

Abstract: Due to the strong privacy of data and the difficulty of annotation, the amount of medical image data is relatively small, which further affects the effect of deep learning. Using data enhancement technology to expand existing data sets can significantly alleviate the problem of insufficient data. Deep Convolutional generative adversarial Network (DCGAN) technology has been widely used in the field of medical image data enhancement, but there are still some problems such as unstable training, difficulty in convergence, easy to produce mode collapse and insufficient quality of the generated images. In this paper, the mode collapse problem of deep convolutional generative adversarial network (DCGAN) is improved by using the fusion attention mechanism and modifying the loss function. Experiments show that the method proposed in this paper can effectively improve the enhancement effect of lung nodule image data.

Keywords: Data enhancement, Pulmonary nodules, Generation of adversarial networks.

1. Introduction

At present, the interpretation of CT images of pulmonary nodules is mainly done by doctors. However, due to the different personal experience of doctors and the fatigue of doctors caused by a large number of film reading, the interpretation of CT images has some problems such as low efficiency and uneven accuracy level. It is urgent to interpret CT images with computer-aided diagnostic CAD.

computer Aided Diagnosis (CAD) is of great help to the reading of lung CT images, which can relieve the pressure of doctors to read the images and guarantee the accuracy to some extent. At present, deep learning technology has been widely used in pulmonary CAD systems, but a large amount of labeled data is required [1]. Due to the strong privacy of patients and the lack of professional doctors to label the data, there are few datasets available for deep learning in medical images. Transfer learning and data enhancement are two commonly used methods to solve the shortage of training data sets [2]. Generative adversarial network (GAN) has excellent image generation capability, so GAN as a data enhancement method has been applied in biomedical image analysis applications [3]. However, GAN image data enhancement technology has problems such as unstable model convergence and low quality of generated images [4].

In this paper, the improved GAN model is used to generate lung nodule images to achieve the purpose of enhancing lung nodule data. The main work of this paper is as follows:

(1) By improving the loss function, using W distance instead of JS divergence and adding regularization terms to the loss function, this paper effectively improves the pattern collapse problem that is prone to occur in DCGAN generated images.

(2) Add the attention mechanism to pay attention to the generation of small nodules to further improve the diversity of generated images.

2. Improved Generative Adversarial Network

2.1. Generative adversarial network

(1) Generate adversarial network introduction

Generative adversarial network (GAN) [5] is a probabilistic generative model proposed by GoodFellow et al in 2014, inspired by minimax two-player games. This model uses the idea of zero-sum game in game theory [6]. GAN is divided into two parts, namely generator part and discriminator part. The input to the generator is random noise, and the input is mapped to a hypothetical data through the generator to generate sample data, and if the generator is good enough, all the generators will correspond to a very real data. The function of the discriminator is to determine the authenticity of the generated sample data relative to the real sample data. During the training process, the discriminator and generator take turns to upgrade, competing with each other to continuously optimize the capabilities of the generator. When the discriminator cannot identify the true data, it means that the generator has generated data that is closer to the real data. Loss function is the core of GAN design [7], and its expression is as follows:

$$L(G,D)=\sum E(\log D_d)+\sum E[\log(1-D(G(p)))] \quad (1)$$

Where, G represents the generator, D represents the discriminator, d represents the real data, and Dd represents the discriminant result of the discriminator on the real data. The goal is to generate data that is consistent with real data.

(2) Disadvantages of generating adversarial networks

GAN is a generative model. Compared with other generative models, which only use backpropagation instead of complex Markov chains, it can generate clearer and more authentic samples. Moreover, it adopts unsupervised learning training, so it has been widely used in medical image enhancement. However, in terms of medical image enhancement, GAN still has problems of unstable training and difficult convergence [8].

2.2. Improved generative adversarial network

Deep Convolutional generative adversarial network (DCGAN) uses deep convolutional neural network in both its generative model and discriminant model, and it is also the first network that combines GAN and convolutional network [2]. The generator for DCGAN is as follows:

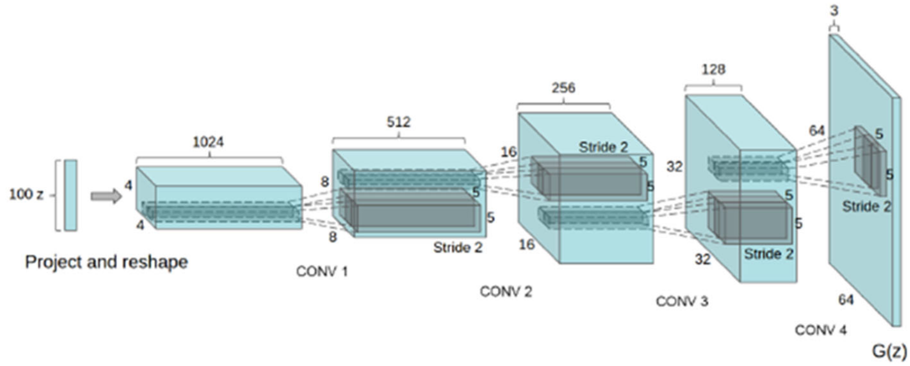

Figure 1. DCGAN model generator

DCGAN has improved the pooling layer, activation function and other aspects to solve the problem of GAN training instability, but there are still problems that the generated images are not diversified enough, that is, mode collapse, and small nodules are difficult to identify. Therefore, the improved generative adversarial network proposed in this paper is mainly improved from the following aspects.

2.2.1. Loss function improvement

Like GAN, DCGAN is composed of generator G and discriminator D. The loss function of the generator model in DCGAN is as follows:

$$\min_G V(D,G) = E_{z \sim pg(z)}[\lg(1-D(G(z)))] \quad (2)$$

The goal of the generator is to generate data as close to the real data as possible, so the goal of the loss function is to get as close to 1 as possible.

In the discriminator model, the loss function is as follows:

$$\max_D V(D,G) = E_{x \sim pdata(x)}[\lg(D(x))] + E_{z \sim pg(z)}[\lg(1-D(G(Z)))] \quad (3)$$

In the above loss function expression, the former expression represents the real data set, so the goal is close to 1; The latter formula is the generated data set, and the discriminator wants to recognize it completely, so the goal of the latter formula is close to 0, and the final goal of the whole formula is close to 1.

Since DCGAN still has the problem of mode collapse, this paper will modify the model from the aspect of loss function. Specific modifications are as follows:

(1) The KL divergence used in DCGAN is asymmetrical, which means that for the same distance, different observation methods will lead to different loss values, resulting in the overall loss decline direction tending to a specific direction, which is easy to cause the problem of model collapse. In this paper, w distance is used to replace KL divergence. The W distance is defined as follows:

$$W[p,q] = \inf_{r \in \prod[p,q]} \iint r(x,y)d(x,y)dxdy = \inf_{r \in \prod[p,q]} E_{(x,y) \sim r(x,y)}[d(x,y)] \quad (4)$$

(2) Add the variance regularization term to the generator loss function, and the formula is as follows:

$$g_{loss} = E_{z \sim px(z)}[\log(1-D(G(z)))] + \lambda\delta(dis(G(z)^i, G(z)^j)) \quad (5)$$

Where gloss represents the generator loss,G(Z) represents the generated image, D (G(Z)) represents the probability of the discriminator on the image G(Z), dis(G(Z)i,G(Z)j) represents the distance between image i and image j in the generated batch of images, & is the variance, and Nimda is the regularization coefficient.

2.2.2. Attentional mechanism

The essence of the attention mechanism is to locate the interesting information and suppress the useless information. In lung nodule images, not all areas in the image contribute equally to the generation of new data. Only the small proportion of lung nodule areas should be concerned, so important parts should be found for processing. Therefore, this paper adds the spatial attention mechanism to make the important part of the generative network more prominent. STN[9] (Spatial Transformer Network) is a model based on spatial attention mechanism, and its structure is as follows:

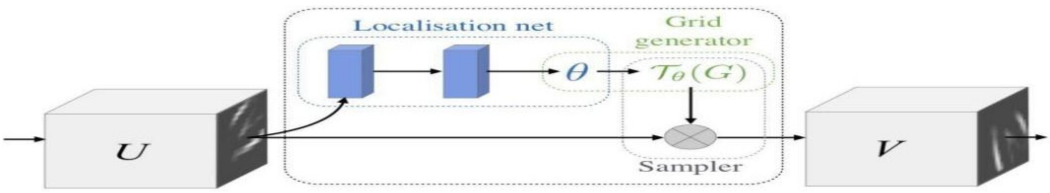

Figure 2. STN model structure diagram

The addition of attention mechanism in DCGAN helps the model to generate micro-nodules.

3. Experiment

3.1. Data set introduction

The data set used in this paper is the LUNA16 dataset, which contains 888 low-dose CT images. Each image contains a series of axial sections of the chest and is annotated by several experienced radiologists. This dataset records the position information and diameter size of lung nodules.

3.2. Experimental environment and related parameters

This experiment was conducted under windows10 platform, the experimental environment was Pytorch1.9.0 +CUDA 12.1, the Graphics Unit (GPU) used was NVIDIA GeForce MX450, and the memory size was 16GB. The 886 images in the LUNA16 data set were randomly scattered, and the training set and test set were divided according to 7:3, that is, 620 pictures were taken as the training set and 266 pictures were taken as the test set. The size of the input picture was 416*416, 500 rounds of training were conducted, and the initial learning rate was 10-3. The evaluation indicators used in this paper are confidence and the initial distance (FID) between the two enhancement frames to generate lung nodule images. The greater the confidence, the greater the possibility that ROI is a nodule. FID represents the difference between the distribution of the real image and the generated image, and represents the distance of the feature vector between the generated image and the real image, and the smaller the distance, the closer the generated nodule is to the real image. The calculation formula for FID is as follows:

$$FID(x,g) = \| \mu_X - \mu_g \|_2^2 + Tr(\sum X - \sum g - 2(\sum X \sum g)^{0.5}) \quad (6)$$

3.3. Experimental result

The original training data, the image generated by unmodified DCGAN and the image generated after modifying the KL divergence in DCGAN to W distance are shown as follows:

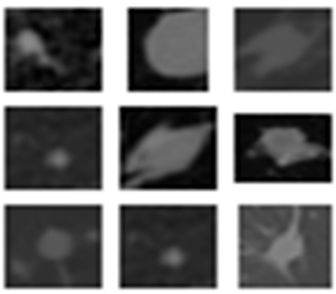

Figure 3. Improved DCGAN to generate images

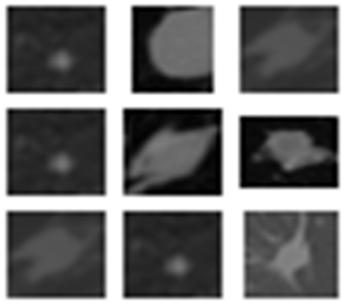

Figure 4. DCGAN generates images

As can be seen from the above pictures, compared with DCGAN that has not been improved, the improved model generates a higher diversity of images, and the model has improved the diversity of generated data to a certain extent.

The FID data results are as follows:

Table 1. Comparison of FID of different frameworks

frame	Data set	FID
DCGAN	LUNA16	85.6
Improved DCGAN	LUNA16	81.3

From the above, it can be seen that the improved DCGAN can improve the diversity of nodules, and reduce the FID value from 85.6 to 81.3, generating more realistic pulmonary nodule data.

4. Conclusion

To solve the problem of enhancement of lung nodule data set, the DCGAN model is improved in this paper. By combining the loss function and the addition of attention mechanism, the problems of pattern collapse in the enhancement of pulmonary nodule data set are improved. The experimental results show that after modification, the model has been improved in the aspects of picture diversity, confidence and FID. However, the disadvantage is that the calculation is relatively complicated. In the future, a relatively simple algorithm improvement method will be studied on the basis of ensuring the diversity of the generated images.

References

[1] LI Yang, GAO Shiqi. Lung Nodule Detection System Based on Data Augmentation and Attention Mechanism[J]. Journal of Beijing University of Posts and Telecommunications, 2022, 45(4): 25-30.

[2] Qi Yongjun , Gu Junhua, Li Weixun et al. Pulmonary nodule image data enhancement Method based on Generative adversarial network [C]// Network Application Branch of China Computer Users Association. Proceedings of the 23rd Annual Conference on New Network Technologies and Applications, China Computer Users Association Network Application Branch, 2019. Editorial Board of Computer Science Science), 2019-5. DOI: 10.26914 / Arthur c. nkihy. 2019.004458.

[3] Kazeminia S, Baur C, Kuijper A, et al. GANs for medical image analysis[J]. Artificial Intelligence in Medicine, 2020, 109: 101938.

[4] Fan Li. Research and application of image data enhancement technology based on generative adversarial network [D].

Zhejiang university, 2022. DOI: 10.27461 /, dc nki. GZJDX. 2022.001690.

[5] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [J]. Advances in Neural Information Processing Systems, 2014, 27

[6] GONG M, NIU X, ZHANG P, et al. Generative adversarial networks for change detection in multispectral imagery [J]. IEEE geoscience and remote sensing letters, 2017, 14 (12); 2310-2314.

[7] Zhang Bin, et al." Overview of Generative Adversarial network improvement Perspectives and Applications." Computer Application Research... Doi: 10.19734 / j.i SSN. 1001-3695.2022.08.0410.

[8] Yang Yu, Based on generating method of image data enhancement against network research [D]. Strategic support troops information engineering university, 2022. The DOI: 10.27188 /, dc nki. Gzjxu. 2022.000039.

[9] Jaderberg, Max, Karen Simonyan, and Andrew Zisserman. "Spatial transformer networks." Advances in neural information processing systems 28 (2015).