Optimization of cardiac magnetic resonance synthetic image based on simulated generative adversarial network

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1155/2021/3279563

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Research Article

Optimization of Cardiac Magnetic Resonance Synthetic Image Based on Simulated Generative Adversarial Network

Ying Fu [1,2] **MinXue Gong** [1] **Guang Yang** [1] **JinRong Hu** [1,2] **Hong Wei** [3] **and Jiliu Zhou** [1,2]

^{1}*School of Computer Science, Chengdu University of Information and Technology, Chengdu 610225, China*
^{2}*Image and Spatial Information 2011 Collaborative Innovation Center of Sichuan Province, Chengdu 610225, China*
^{3}*Department of Computer Science, University of Reading, Reading RG6 6AY, UK*

Correspondence should be addressed to JinRong Hu; hjr@cuit.edu.cn

The generative adversarial network (GAN) has advantage to fit data distribution, so it can achieve data augmentation by fitting the real distribution and synthesizing additional training data. In this way, the deep convolution model can also be well trained in the case of using a small sample medical image data set. However, some certain gaps still exist between synthetic images and real images. In order to further narrow those gaps, this paper proposed a method that applies SimGAN on cardiac magnetic resonance synthetic image optimization task. Meanwhile, the improved residual structure is used to deepen the network structure to improve the performance of the optimizer. Lastly, the experiments will show the good result of our data augmentation method based on GAN.

1. Introduction

The deep convolutional network model [1] has a large number of parameters and needs a large number of labelled data sets to train for some computer vision tasks such as image classification [2, 3] and image segmentation [4]. Insufficient data sets or unbalanced data sets will lead to overfitting of the deep convolutional model [5]. Therefore, the practical application of the deep convolution model is not wide, which requires a large medical data set to fully train the auxiliary diagnostic model. Insufficient data not only fails to have the effect of auxiliary diagnosis but also causes interference to the doctor's diagnosis [6].

The requirement of medical data needs expensive equipment and experienced doctors to annotate the data, which is extremely time-consuming. In addition, the privacy of patients in the process of collecting medical image data is quite sensitive, and most of patients are reluctant to contribute their own data. Such factors make it difficult to obtain large-scale labelled medical image data sets. Although there are many public medical data sets available on the Internet,

most of medical data sets are of limited size and are only applicable for specific medical problems. Deep convolution model cannot be fully trained with insufficient data sets and cannot be truly used in clinical diagnostic tasks due to the safety of medical diagnosis.

Data augmentation is an implicit regularization method for data, which solves the scarcity of data by manually extending the training set. Recently, a method proposed by Krizhevsky [7] to flip and rotate images in the training set is widely used in data augmentation. Therefore, the performance of deep convolution model is improved because that enables the model to recognize the image invariance that is too small to be seen by the naked eyes.

However, conventional data augmentation methods can only produce transform data with limited features based on raw data and do not essentially add new data to the original data set, so many scholars have proposed data augmentation methods based on GAN [8]. GAN-based data augmentation method has been successfully applied in the field of medical imaging. Ali et al. [9] proposed a method to augment skin injury image data based on progressive GAN (PGGAN) and

improved the large-scale feature dependence of PGGAN by using self-attention mechanism, which also effectively improved the classification accuracy of skin injury data. Frid-Adar et al. [10] applied DCGAN [11] and ACGAN [12] to liver injury data, and the experimental results showed that DCGAN had a stronger data augmentation effect than ACGAN in the case of fewer data sets.

2. Data Augmentation Method Based on E-GAN

GAN is a generative model proposed by Ian Goodfellow et al. [8], which is composed of the generator G and the discriminator D. The generator G uses the noise z sampled in the uniform or normal distribution as input synthetic image $G(z)$. The discriminator D tries to discriminate the synthetic image $G(z)$ as false and the real image x as true. The parameters of each model are adjusted by successive adversarial training. Finally, the generator obtains the real sample distribution model and obtains the performance to generate near-real images.

The training process of GAN aims to find the balance between the generator and the discriminator, which can be expressed in the following equation:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim P_r(x)}[\log D(x)] \\ + \mathbb{E}_{z \sim P_z(z)}[\log(1 - D(G(z)))],$$

(1)

where P_r represents the real sample distribution, P_g represents the generated sample distribution, $D(x)$ is the discriminator determining whether the real image of the input comes from the real data distribution or the probability of the synthetic distribution, and $D(G(z))$ represents that the discriminator thinks the synthetic image comes from the synthetic distribution or the real distribution.

The generator synthesizes new samples by fitting the original sample distribution. The new samples are synthetic from the distribution learned from the generative model, which makes them have new features different from the original sample. This property makes it possible to use the synthetic sample as a new training sample to achieve data augmentation. Considering the improvement of the diversity performance of the generator by evolutionary GAN (E-GAN) [13], we apply it to the augmentation of cardiac magnetic resonance synthetic image.

Firstly, to ensure the consistency of the training data, all samples are normalized in this experiment. Secondly, the GAN training set is enhanced by horizontal flip, vertical flip, and vertical and horizontal axis 0–2% random magnification translation before training to avoid the loss of image information. Thirdly, the new images are synthetic when the GAN training is completed. Lastly, the synthetic and original images as the training set are used for the classification network.

The architecture of the GAN model based on E-GAN is shown in Figure 1. The model consists of two parts: mutation and adaptive score evaluation. After these two processes, the scores are sorted and then the generator with the highest current scores will be selected.

In E-GAN, the father generators are derived by using different mutational methods. These mutational operators are actually different training goals to reduce the distance between the synthetic distribution and the real data distribution through different angles. Meanwhile, the best discriminator should be trained before every mutation operation D*. Then the optimal generator is selected in E-GAN in the current environment through three mutation methods and adaptive score scores.

Different from the original E-GAN model, the input of our discriminator is mixed up [14] from two original samples interpolated into one sample, and the actual label of this sample is also converted to label by interpolation method. Therefore, the minimization goal of the model becomes the cross-entropy loss between the predicted labels and the interpolated samples of the discriminator. The training process can be broadly described as the following two repetitive processes:

(1) The generator stops training until the current optimal generator is found in the way of evolutionary algorithms, so the sample is synthetic with this generator. Meanwhile, the synthetic sample is interpolated linearly with the real sample to produce new sample and corresponding interpolation label. The interpolation sample and the interpolation label are used as the input of the discriminator.

(2) The discriminator is trained with interpolated sample and interpolated label, and the discriminator parameter value is fixed after the parameter is updated. This discriminator with stronger discriminative ability becomes the competitive environment of the next-round generator survival of the fittest; thus the generator parameters are updated more effectively.

However, the peak signal-to-noise ratio (PSNR) [15] and structural similarity (SSIM) [16] indexes commonly used in medical image restoration cannot be applied to the performance evaluation of data augmentation methods because they are hard to be used to evaluate the diversity of synthetic samples. Therefore, in this paper, binary classification of benign and malignant medical image as the downstream task is chosen to verify the effectiveness of our method. In this paper, two different classifiers are used in the downstream classification task, ResNet-50 [17] as classifier 1 and Xception [18] as classifier 2, which are often used in medical image classification tasks.

2.1. Synthetic Images Optimization Method Based on SimGAN. However, there is still a large gap between the synthetic distribution from GAN and the real distribution. To narrow this gap, Ashish et al. proposed a simulated generation adversarial network (simulated GAN, SimGAN) [19].

Different from the conventional GAN model, SimGAN sets an image optimizer (refiner) between the generator and the discriminator, which replaces the role of the generator in the GAN and is the ultimate training object of the SimGAN.
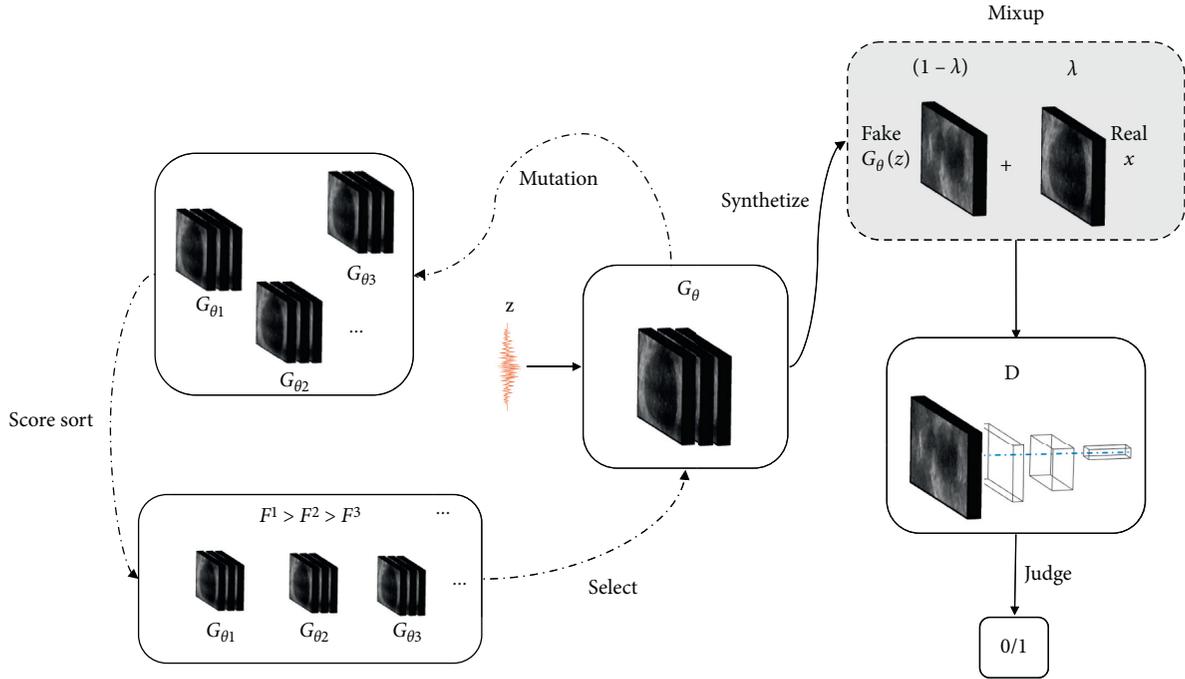
FIGURE 1: E-GAN Architecture with mixup.

The input of the optimizer is a synthetic image rather than a noise variable to generate optimized image, which will be input into the discriminator. Simultaneous and adversarial training for discriminator and optimizer can achieve optimized performance. The specific structure of SimGAN is shown in Figure 2.

On one hand, the optimizer needs to reduce the distance between the synthetic image distribution and the real image distribution. On the other hand, the discriminator needs to classify the real and optimized image correctly as far as possible. The discriminator network updates the parameters by minimizing the loss function shown in the following equation:

$$L_D(\phi) = -\sum_i \log\big(D_\phi(\tilde{x}_i)\big) - \sum_j \log\big(1 - D_\phi(y_j)\big), \quad (2)$$

where $\tilde{x}_i$ represents the optimized image of the ith synthetic image, y_j represents the yth truth image of the first sheet, D_φ represents the discriminator, and φ is the parameter of the discriminator network. The formula can be equal-price discriminator D_φ for the cross-entropy loss of the dichotomous problem, slightly different from the original GAN. The final output of the discriminator D_φ is the probability value of the optimized image. When training SimGAN, each small batch training data is composed of optimized image and real image randomly. The goal of training is to judge the real image y_j as far as possible as 0 and the optimized image $\tilde{x}_i$ as far as possible as 1. Finally, the parameters φ are updated by random gradient descent method.

Notably, SimGAN is used to prevent the synthetic image from getting too close to the real image and thus completely losing the image features of the synthetic image itself; L1 regular term [20] is set as a self-regular term to retain some important semantic information in the synthetic image. The loss function that needs to be minimized during the training of the optimizer is shown in the following equation:

$$L_R(\theta) = -\sum_i \log\big(1 - D_\phi(R_\theta(x_i))\big) + \lambda\big\|\psi(R_\theta(x_i)) - \psi(x_i)\big\|_1,$$

$$(3)$$

where ψ represents the identity mapping method, R_θ represents the optimizer, θ represents the internal parameters of the optimizer, and $\|.\|_1$ represents the L1 regular term.

In order to make the medical synthetic image closer to the real image and further improve the classification accuracy, this paper designs an optimization method of medical synthetic image based on SimGAN. The flow chart of the medical synthetic image optimization method is shown in Figure 3.

This method can be divided into three experimental steps:

(1) Firstly, we use the trained E-GAN model to synthesize each kind of image of each medical data set, add the E-GAN synthetic medical image to the original training set to expand the training set, and use the expanded training set as the new training set to train classifier 1.

(2) Secondly, we use the expanded training set to train SimGAN, and then we optimize the composite image. After that, we mix the optimized image and
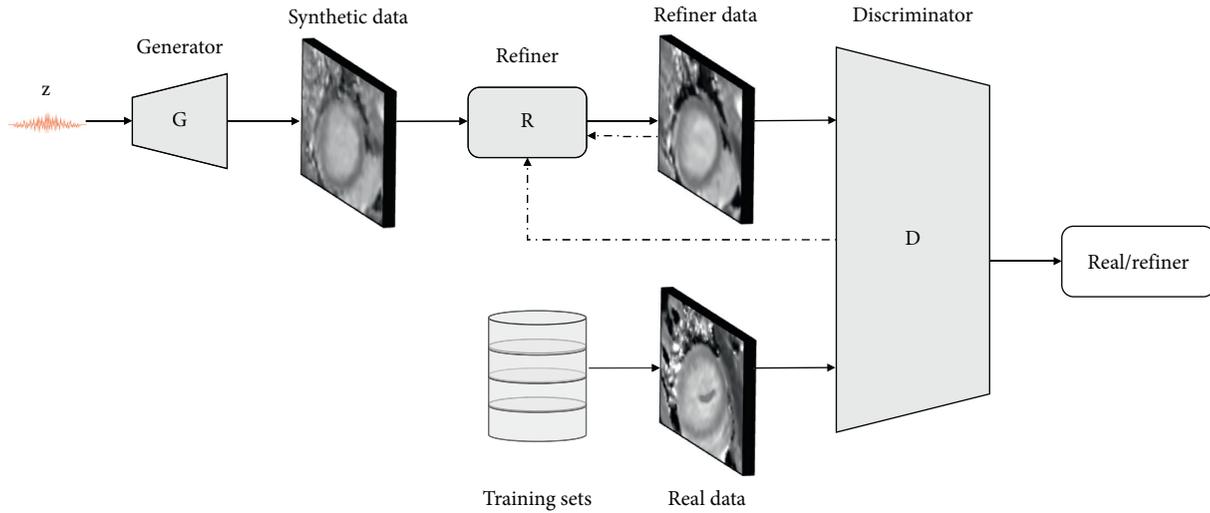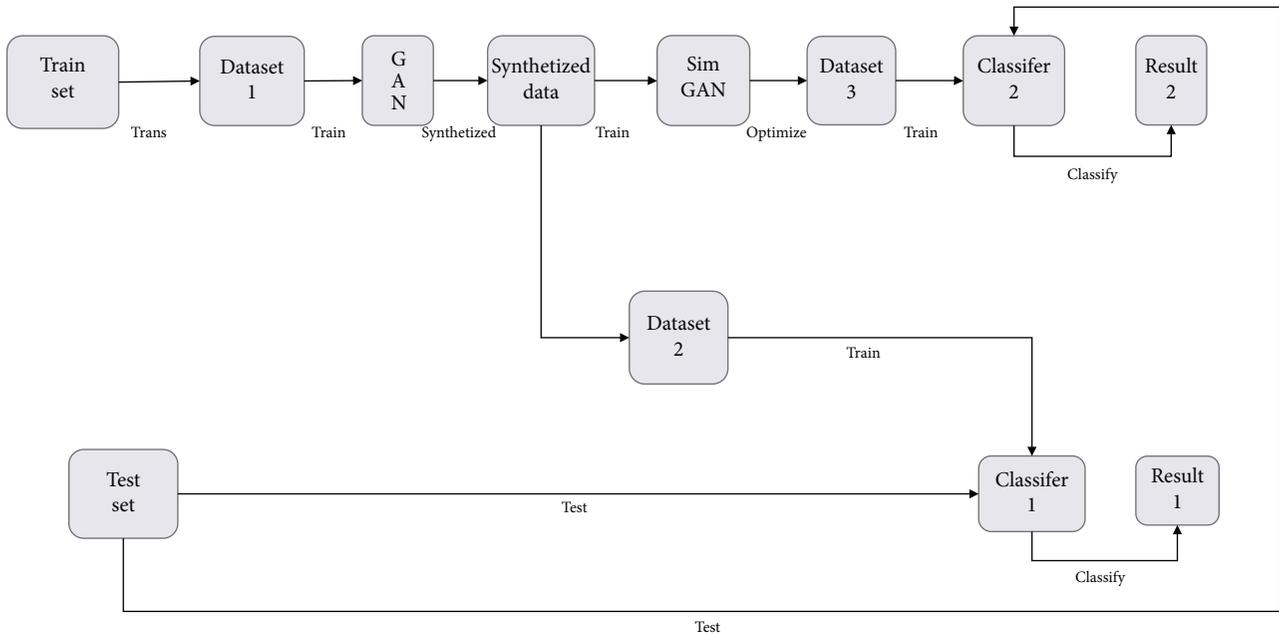
Figure 2: SimGAN architecture.

Figure 3: The flow chart of the medical synthetic image optimization.

the original training set to produce the optimized augmented training set and use the training set to train classifier 2.

(3) Finally, the test set is input into classifier 1 and classifier 2, and the test results are compared to obtain the experimental results of this method.

Unlike the internal structure of the GAN and its conventional variant models, in order to make the discriminator extract more information and further increase the depth of the model, so that the network can retain more feature information, the SimGAN discriminator uses a full-volume network instead of a network structure with a full-connected layer, so that the output of each layer can retain the feature information of the previous layer as much as possible. The

SimGAN optimizer is very different from the GAN generator structure. Its output is an optimized image with the same size as the input synthetic image and does not require upsampling, so there is no need to use a transposed convolutional layer with a step size of 2. The optimizer is designed as a residual network that allows the optimizer to retain the global structure of the synthetic image rather than completely changing the content of the synthetic image.

The computation of the SimGAN is relatively small. In order to further improve the fitting ability of the model, this experiment is based on the residual structure, and we use PReLU to replace the ReLU activation function; besides, the original residual structure will reenter the added result ReLU activation function.

However, the original residual structure [17] has been improved in some recent experiments. For example, the article in [21] deleted the final ReLU [22] activation function of the original residual structure, which makes the model performance improved to some extent. Therefore, the ReLU activation function is not added to the tail of the residual structure in this paper. The residual structure does not produce additional parameters. Replacing the original convolution layer with the residual structure does not augment the computational complexity. The model can still be smoothly back-propagated to achieve end-to-end training. Combined with the residual structure shown in PReLU [23], the residual structure used in this experiment is shown in Figure 4.

In the experiment, five residual structures as shown in Figure 4 are set for the optimizer as the middle layer, which improves the depth of the optimizer and makes it have stronger fitting performance. Similarly, in order to increase the ability of the discriminator, two residual structures are added to the discriminator. The internal construction of the model of the optimizer and discriminator is shown in Table 1.

The structure of the SimGAN is relatively simple and clear, and, compared with the DCGAN structure, it is not very different and does not need to debug for a long time, but in the process of code implementation there are two noteworthy problems:

(1) With the advance of training, the optimized image of the optimizer is constantly changing, but the image that the discriminator can accurately discriminate will also be concentrated in the recent optimized image range, which will lead to the model training not converging. Because the discriminator forgets the previously learned features, the optimizer reintroduces the global image features and produces artifacts again. To solve this problem, a buffer needs to be set to store the optimized image generated during training. The buffer area structure adopted in this experiment is shown in Figure 5.

As shown in Figure 5, in order to introduce global image features, in the process of training the discriminator of the SimGAN, whenever the discriminator needs a small batch of input images, the discriminator extracts half of the optimized image from the buffer and then the other half of the optimized image from the output of the optimizer. After this round of training, use the remaining half of the current optimizer to optimize the image to fill the position of the vacancy in the buffer.

(2) SimGAN training needs to be divided into two stages, which are only automatic setting of parameters from the point of view of code implementation. The first stage is the pretraining process. First, it is necessary to fix the parameters of the discriminator and let the optimizer train 500
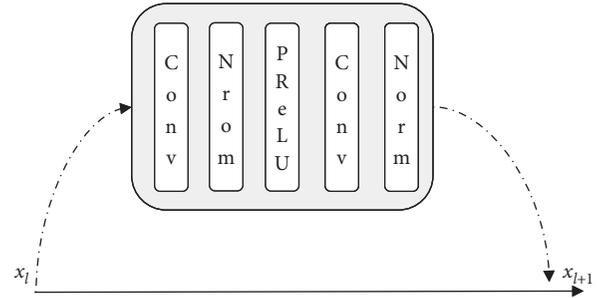

Figure 4: The residual structure used in this paper.

Table 1: The internal construction of the SimGAN used in this paper.

Refiner	Discriminator
Input image	Input image
Conv ([3 × 3], stride = 1)	Conv ([3 × 3], stride = 2)
Residual structure ([3 × 3] × 2)	Conv ([3 × 3], stride = 2)
Residual structure ([3 × 3] × 2)	Max pooling (stride = 1, 3 × 3)
Residual structure ([3 × 3] × 2)	Residual structure ([3 × 3] × 2)
Residual structure ([3 × 3] × 2)	Residual structure ([3 × 3] × 2)
Residual structure ([3 × 3] × 2)	Conv ([1 × 1], stride = 1)
Conv ([1 × 1], stride = 1)	Conv ([1 × 1], stride = 1)
Output image	Softmax

rounds continuously. Then, it is necessary to fix the parameters of the optimizer and let the discriminator train 100 rounds continuously. The second stage is the formal training process. The optimizer and discriminator update their own parameters with a training frequency of $2:1$.

The training process of SimGAN can be described as follows:

(1) We will first sample a batch-processed composite image from the E-GAN composite image, get a batch-processed optimized image from the synthetic image input optimizer, calculate the average loss in the optimized image input discriminator, and then update the parameters of the optimizer using gradient descent method (Stochastic Gradient Descent, SGD) [24]. After the optimizer updates several parameters, the parameters of the optimizer are fixed into the training stage of the discriminator.

(2) In the training stage of the discriminator, we first need to sample a batch image from the training set of the synthetic image and the original image and put the synthetic image into the optimizer to get a batch optimization image, sample half batch optimization image from the buffer area, and extract the other half optimization image from the optimizer output, input the optimized image with the real image into the discriminator, calculate the discriminator average loss, and update its parameters by minimizing the discriminator loss by gradient descent method.
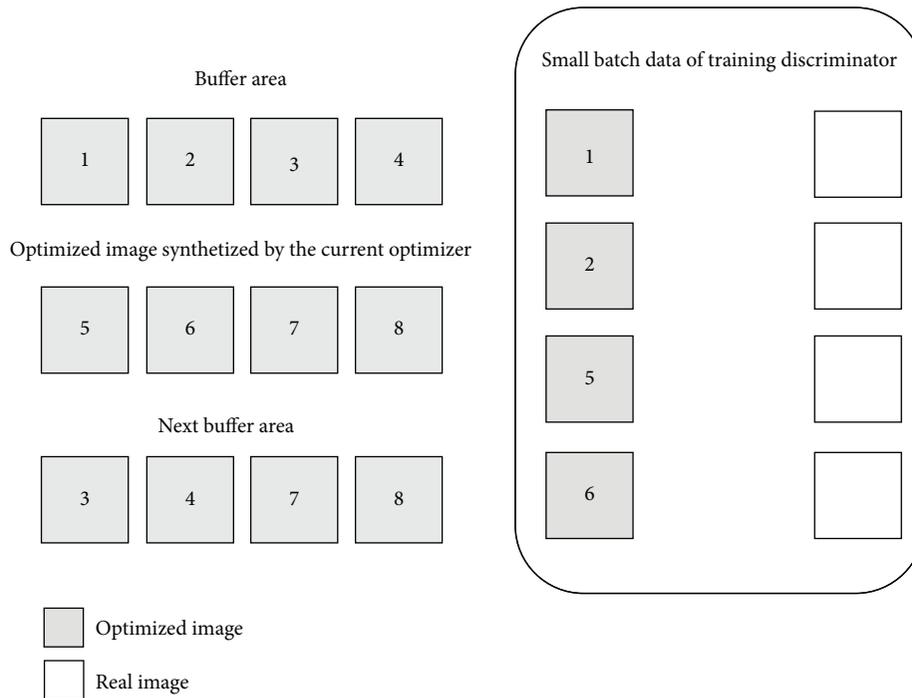
Buffer area

Optimized image synthetized by the current optimizer

Next buffer area

Small batch data of training discriminator

□ Optimized image

□ Real image

FIGURE 5: The buffer area structure.

(3) After this round of discriminator training process is completed, the remaining optimized images output by the current optimizer are used to replace the vacancy position in the buffer.

3. Results and Discussion

For the purpose of verifying the effectiveness of this study, two different medical image data sets are used as the application objects of the data augmentation method. After k-fold dynamic partition data, four training setsare used to train two different classifiers after amplifying the training set. Finally, the average classification results of the test set corresponding to the four training sets are used to verify the concrete effect of the proposed data augmentation method.

3.1. Data set. Cardiac magnetic resonance imaging technology [25] is a noninvasive means of examination. It has a relatively large field of vision and more parameters. Its imaging has the characteristics of multisequence and multiplane.

Because it only needs one imaging to show the whole structure of the heart and myocardial function, cardiac magnetic resonance imaging technology is widely used in clinical diagnosis and evaluation, which is considered to be the most comprehensive and accurate noninvasive examination method. The cardiovascular magnetic resonance data from this experiment were from cooperative hospitals. All samples were 2d short-axis native t1 mapping magnetic resonance images. The spatial spacing of these cardiac magnetic resonance images ranged from

$1.172 \times 1.172 \times 1.0$ to $1.46 \times 1.46 \times 1.0$ mm3. The original pixel size was $256 \times 218 \times 1$. The image of benign and malignant marking and segmentation of the region have senior experts manual marking and drawing. The original image data is in ".mha" format; 298 images were obtained by resampling, region of interest selection, normalization, and other preprocessing processes, 221 myocardial patient images and 77 nonpatient images, respectively. The preprocessed image size was $80 \times 80 \times 1$.

3.2. Experimental Environment. The training of E-GAN requires extremely high experimental equipment, so this experiment is all deployed in Linux servers equipped with multiple Tesla M40 graphics cards. In this paper, TensorFlow [26] is chosen as the deep learning framework to build E-GAN model. The experiment of SimGAN is basically the same as the E-GAN experiment environment, because the SimGAN needs very frequent debugging parameters. Therefore, Keras [27] is selected as deep learning framework that can adjust the parameters conveniently and realize the network quickly.

3.3. Experimental Results and Analysis of E-GAN. After the training of E-GAN, the comparison results of the real disease samples, the nonpatient samples, and the disease samples synthetic by the generator and the nonpatient samples of the two data sets are shown in Figure 6.

From Figure 6, it can be found that the synthetic image observed by the human eye is very similar to the real image, and there is not much difference in the first perception. But it can be seen from careful observation that the sharpness of
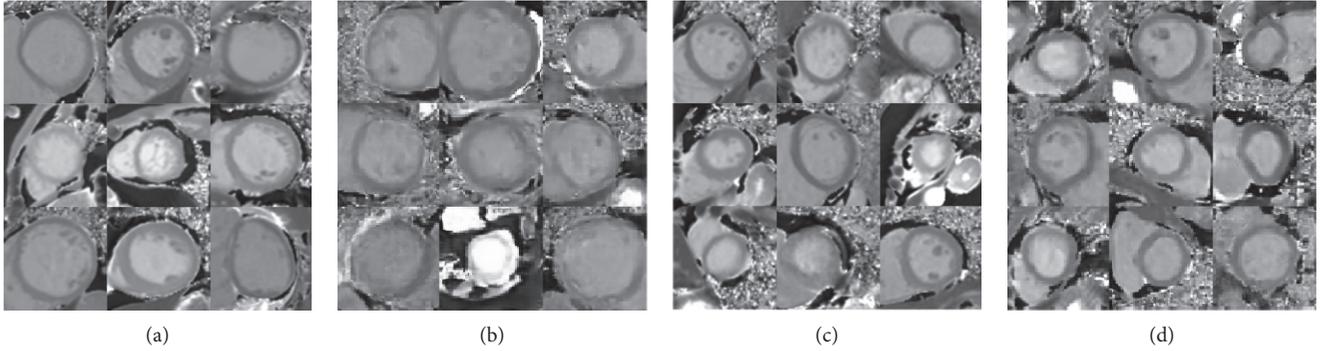
FIGURE 6: Real and synthetic images of cardiac magnetic resonance images. (a) Real disease images. (b) Synthetic disease images. (c) Real nonpathological images. (d) Synthetic nonpathogenic images.

TABLE 2: Average classification results of cardiac images.

Classifier	Origin (%)	Classic Aug.	E-GAN
ResNet-50	77.7	80.9	84.8
Xception	79.5	82.8	86.9

the composite image is slightly worse than that of the real image, the contour is not sharp in the real image, and the edge part is also relatively blurred.

The use of observation method is highly subjective, so observation method can only be used as a reference evaluation standard. The main purpose of this paper is to augment the data of medical images. Therefore, two expanded training sets are used to train ResNet-50 and Xception classifiers, respectively, and compared with two classifiers trained using the original training set, so as to verify the E-GAN data augmentation effect. In order to evaluate the effect of the data augmentation method proposed in this paper more objectively, average classification results of cardiac images using ResNet-50 and Xception are shown in Table 2.

The results of ResNet-50 classification show that the accuracy of the test is improved by 2.7% using only the single-sample space geometric transformation data, and the average classification accuracy of the E-GAN is 84.8%. The experimental results in Xception classifier are shown to be consistent with most of ResNet-50, and the classification results of Xception classifier are slightly better, which is related to the fitting performance of the model itself. The average classification accuracy result is 86.9%.

The training effect of adding more data volume classifier should be better in theory, but this experiment found that the more synthetic images added to the original training set is not the more the better, and the classification effect does not rise and fall after adding a certain number of synthetic images. The above classification results were obtained when using E-GAN to synthesize 3-fold cardiac magnetic resonance images and adding the original training set. During the E-GAN experiment, the classification accuracy of cardiac magnetic resonance images changed with the proportion of synthetic images in the training set as shown in Figure 7.

The horizontal axis marked as nonsynthetic image indicates the experimental results of nonsynthetic image, and the labelled augmentation and nonaugmentation classifier indicate whether the single-sample spatial geometric transformation data augmentation method is used when training the current classifier, respectively.

The trend of Xception classification results is roughly the same as that of ResNet-50, and it reaches the best after adding 3 times synthetic image, reaching 86.9%. This is because there is still a certain gap between the synthetic image and the real image, adding too many synthetic images will make the proportion of the real image too small, and the classifier overfits the synthetic image.

3.4. Experimental Results and Analysis of SimGAN. After obtaining the above experimental results, we have carried out the optimization experiment of synthetic images. During the SimGAN training process, as the optimization performance continues to improve, the changes of the synthetic images of the cardiac magnetic resonance data set are shown in Figure 8.

It can be observed from Figure 8 that the synthetic image of dirty magnetic resonance in the center of the optimization process will not be very different from the original synthetic image, with only some minor changes at a time. However, after continuous training, the optimizer will make the optimized synthetic image have more abundant texture information that is closer to the real image.

After training the SimGAN, this paper uses the optimizer of E-GAN synthetic cardiac magnetic resonance image input SimGAN. Each input can obtain the same number of optimized medical synthetic images. The synthetic images before and after optimization of the two data sets are shown in Figure 9.

It can be observed that the optimized cardiac magnetic resonance synthetic image has more details than the original synthetic image and appears sharper in some marginal parts, and the overall characteristics change a little, but it can be observed from the naked eye to become clearer.

The quality change of the image observed from the naked eye is only a subjective evaluation standard, or it is necessary to prove the final effect of the method through
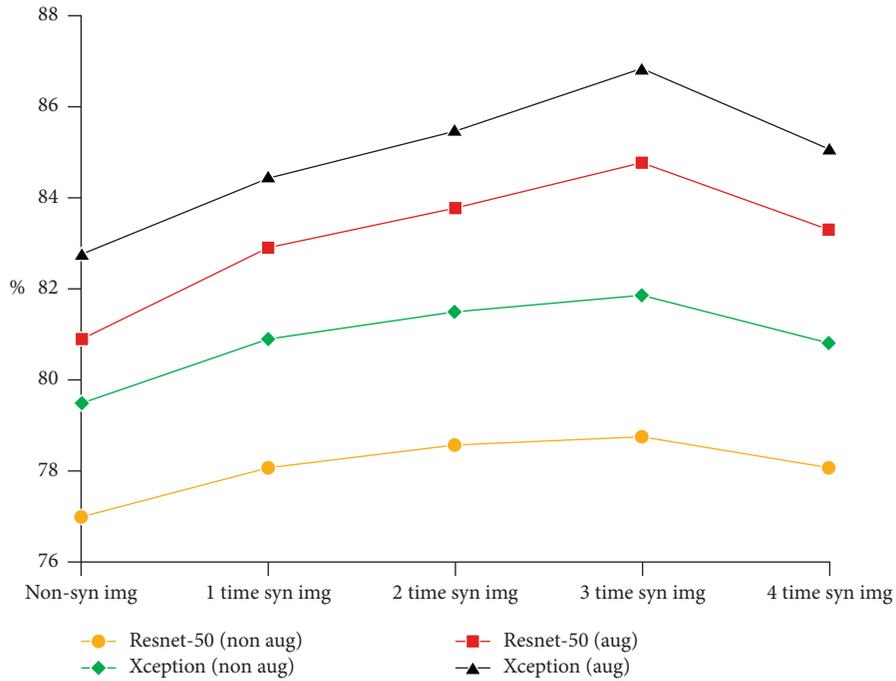
Figure 7: The classification accuracy of cardiac magnetic resonance images changed with the proportion of synthetic images the in training set.
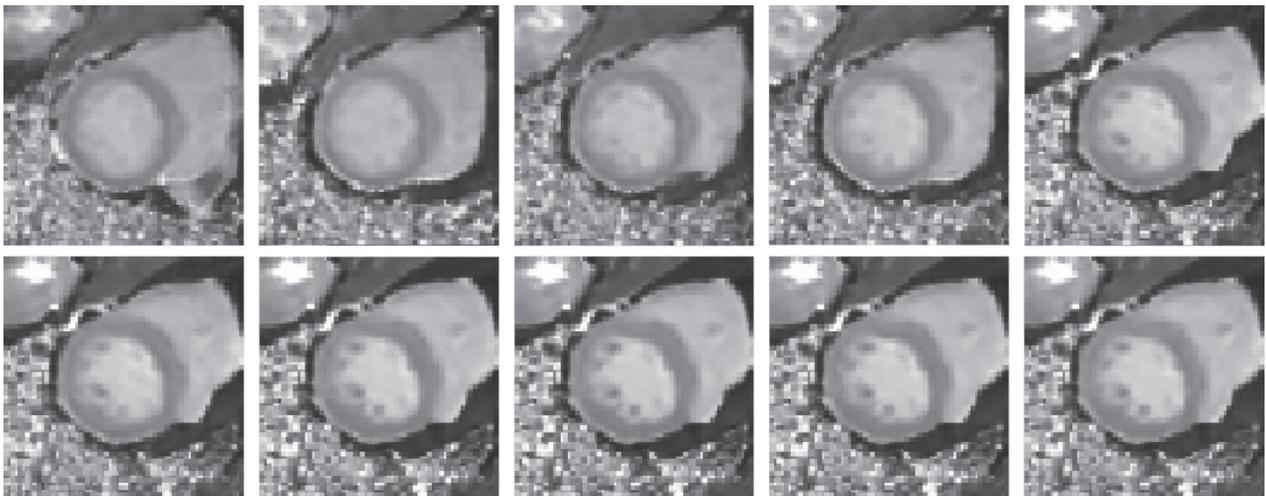

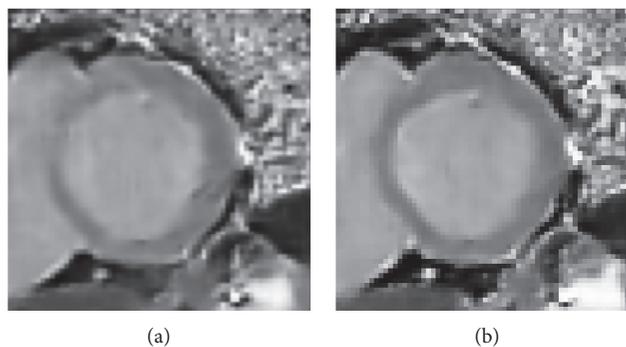Figure 8: The synthetic images in optimization process of the cardiac magnetic resonance data.

Figure 9: Comparison of cardiac magnetic resonance synthetic imaging and optimized images. (a) Synthesized images. (b) Optimized images.

TABLE 3: Average classification results of cardiac images.

Classifier	Non (%)	Classic Aug. (%)	E-GAN (%)	SimGAN (%)
ResNet-50	77.7	80.9	84.8	86.1
Xception	79.5	82.8	86.9	88.0

the classification results. During the training of cardiac magnetic resonance images using 4 times synthetic image, the average classification accuracy is optimal. Compared with E-GAN data augmentation experiments, adding a larger proportion of optimized images can further improve the classification results, which indicates that the optimized synthetic images are closer to the real images than the original synthetic images. The final classification results obtained in this experiment are shown in Table 3.

From Table 3, it can be observed that, in the classification experiment of cardiac magnetic resonance images, the average classification accuracy was increased by 1.3% and 1.2% in the two classifiers after adding 4 times of optimized images, and the final average classification accuracy reached 86.1% and 88.0%, respectively. The average accuracy of this method is improved in the classification experiments of cardiac magnetic resonance imaging, and the optimized composite image is closer to the real image, which enables the classifier to be trained more fully.

4. Conclusions

In order to narrow the gap between synthetic and real images of cardiac magnetic resonance and further optimize the details of synthetic images to augment the effect of GAN-based data augmentation methods, this paper proposes a medical synthetic image optimization method based on SimGAN. The modified residual structure is used to redesign, model, and apply it to medical synthetic images output by E-GAN. The experimental results show that the method of optimizing cardiac magnetic resonance synthesis image based on SimGAN can further improve the classification accuracy and achieve the expected results.

Data Availability

The data used in the work are taken from the cooperative hospital. The hospital requires that the data should not be disclosed and be only used for research.

Conflicts of Interest

The authors declare that they have no conflicts of interest related to this work.

Acknowledgments

References

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[2] J. Zhang, P. Liu, F. Zhang, and Q. Song, "CloudNet: ground-based cloud classification with deep convolutional neural network," *Geophysical Research Letters*, vol. 45, no. 16, p. 8665, 2018.

[3] H. Xiao, F. Zhang, Q. He et al., "Classification of ice crystal habits observed from airborne cloud particle imager by deep transfer learning," *Earth and Space Science*, vol. 6, no. 10, pp. 1877–1886, 2019.

[4] D. Wang, A. Khosla, R. Gargeya et al., "Deep learning for identifying metastatic breast cancer," 2016.

[5] N. Srivastava, G. Hinton, A. Krizhevsky et al., "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[6] F. T. De Dombal, D. J. Leaper, J. R. Staniland, A. P. McCann, and J. C. Horrocks, "Computer-aided diagnosis of acute abdominal pain," *British Medical Journal*, vol. 2, no. 5804, pp. 9–13, 1972.

[7] A. Krizhevsky, "In Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 21, pp. 1097–1105, 2012.

[8] I. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "In Generative adversarial nets," *Advances in Neural Information Processing Systems*, vol. 41, pp. 2672–2680, 2014.

[9] S. A. Ibrahim and M. F. M. Y. B. Mahdy, "Data augmentation for skin lesion using self-attention based progressive generative adversarial network," 2016, https://arxiv.org/abs/1905.11960.

[10] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification," *Neurocomputing*, vol. 321, pp. 321–331, 2018.

[11] A. Radford and L. M. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *Computer Science*, vol. 45, 2015.

[12] X. Xia, R. Togneri, F. Sohel et al., "Auxiliary classifier generative adversarial network with soft labels in imbalanced acoustic event detection," *IEEE Transactions on Multimedia*, vol. 21, no. 6, pp. 1359–1371, 2018.

[13] C. Wang, C. Xu, X. Yao, and D. Tao, "Evolutionary generative adversarial networks," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 6, pp. 921–934, 2019.

[14] H. Zhang, M. Cisse, Y. N. Dauphin et al., "Mixup: beyond empirical risk minimization," 2017.

[15] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electronics Letters*, vol. 44, no. 13, pp. 800-801, 2008.

[16] Z. Wang, "The SSIM index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 13, no. 4, 2014.

[17] K. He, X. Zhang, S. Ren et al., "In deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, July 2016.

[18] F. Chollet, "In xception: deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1251–1258, Honolulu, HI, USA, July 2017.

[19] A. Shrivastava, T. Pfister, O. Tuzel et al., "In Learning from simulated and unsupervised images through adversarial training," in *Proceedings of the IEEE Conference on Computer*

Vision and Pattern Recognition, pp. 2107–2116, Honolulu, HI, USA, July 2017.

[20] M. Y. P. T. Hastie, "L1-regularization path algorithm for generalized linear models," *Journal of the Royal Statistical Society: Series B*, vol. 69, no. 4, pp. 659–677, 2007.

[21] C. Ledig, L. Theis, F. Huszár et al., "In photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4681–4690, Honolulu, HI, USA, July 2017.

[22] X. Glorot and B.Y. Antoine, "In deep sparse rectifier neural networks," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, pp. 315–323, Ft. Lauderdale, FL, USA, April 2011.

[23] K. He, X. Zhang, S. Ren et al., "In delving deep into rectifiers: surpassing human-level performance on imagenet classification," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1026–1034, Santeigo, CL, USA, December 2015.

[24] L. . Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010*, pp. 177–186, Springer, Paris France, August 2010.

[25] A. M. Maceira, J. Joshi, S. K. Prasad et al., "Cardiovascular magnetic resonance in cardiac amyloidosis," *Circulation*, vol. 111, no. 2, pp. 186–193, 2005.

[26] M. Abadi, A. Agarwal, B. Paul et al., "Tensorflow: large-scale machine learning on heterogeneous distributed systems," 2016, https://arxiv.org/abs/1603.04467.

[27] Chollet, F., https://github.com/fchollet/keras, 2015.