Research article

dm-GAN: Distributed multi-latent code inversion enhanced GAN for fast and accurate breast X-ray image automatic generation

Jiajia Jiao[1],*, Xiao Xiao[1] and Zhiyu Li[2],*

[1] College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China
[2] Department of Medical Imaging, Shanghai East Hospital, Tongji University School of Medicine, Shanghai 201306, China

* Correspondence: Email: jiaojiajia@shmtu.edu.cn, lychee71@163.com.

Abstract: Breast cancer seriously threatens women's physical and mental health. Mammography is one of the most effective methods for breast cancer diagnosis via artificial intelligence algorithms to identify diverse breast masses. The popular intelligent diagnosis methods require a large amount of breast images for training. However, collecting and labeling many breast images manually is extremely time consuming and inefficient. In this paper, we propose a distributed multi-latent code inversion enhanced Generative Adversarial Network (dm-GAN) for fast, accurate and automatic breast image generation. The proposed dm-GAN takes advantage of the generator and discriminator of the GAN framework to achieve automatic image generation. The new generator in dm-GAN adopts a multi-latent code inverse mapping method to simplify the data fitting process of GAN generation and improve the accuracy of image generation, while a multi-discriminator structure is used to enhance the discrimination accuracy. The experimental results show that the proposed dm-GAN can automatically generate breast images with higher accuracy, up to a higher 1.84 dB Peak Signal-to-Noise Ratio (PSNR) and lower 5.61% Fréchet Inception Distance (FID), as well as 1.38x faster generation than the state-of-the-art.

Keywords: breast X-ray image; Generative Adversarial Network; multiple latent code inversion; multiple discriminators

1. Introduction

Breast cancer is a phenomenon of uncontrolled proliferation of breast epithelial cells under the action of various carcinogens, and its mortality rate is up to 33% of the total cancer deaths [1]. Breast X-ray photography is an effective method for early breast cancer screening based on the doctor's diagnosis of the different signs in the mammary X-ray image [2], where breast masses are the most obvious pathological signals [3,4]. Due to the advantages of the high-precision of breast X-ray images to improve the detection accuracy, mammography has become one of the most effective methods for breast diagnosis. With the increasingly widespread application of computer-aided medical diagnosis technology [5], breast assisted diagnosis technology based on breast X-ray images applies artificial intelligence algorithms to solve many problems such as mass detection, mass segmentation and mass classification. For example, Redmon et al. proposed a faster Region based Convolutional Neural Network (R-CNN) [6] to achieve tumor target area detection. Long et al. simplified the feature extraction part of the segmentation model based on fully convolutional networks (FCN) [7], effectively improving the performance of tumor segmentation. Arevalo et al. used neural networks for hierarchical learning of tumor features [8] to improve the accuracy of tumor classification. These artificial intelligence algorithms on breast X-ray images can assist the detection and diagnosis of breast cancer [9], simplify the diagnosis process of doctors and improve the accuracy of diagnosis.

In order to improve the accuracy of mass detection, and segmentation, a large number of high-precision breast image data sets are usually required as fundamental support. However, it is challenging to use the traditional data augmentation methods such as medical device collection because of some difficult problems with the long time requirements, differences in collection devices, loss of data information and the efforts of doctors. Therefore, achieving high-precision breast medical image generation can provide valid datasets for medical assisted diagnosis technology, as well as alleviates the pressure of modern medical image data scarcity.

Generative Adversarial Networks (GAN) is a mainstream deep learning framework for image generation. GAN was first proposed by Ian Goodfellow and his team in 2014 [10]. It establishes a game relationship between the generator and discriminator, aiming to optimize the output of the generator to approximate the true data distribution. Subsequently, the enhanced GAN versions such as Wasserstein GAN (WGAN) could address issues of training instability and mode collapse [11]. Subsequently, the Conditional Generative Adversarial Networks (cGANs) allow the generator to generate samples based on specific conditions [12], and further expand the applications of GANs. Since 2017, the application of GAN in the field of medical imaging has significantly increased [13]. As for the breast medical imaging, GANs have also played a significant role in intelligent medical diagnosis. More importantly, GAN [10] and various variant GANs [14–20] have been proposed for medical image augmentations. These enhanced GANs try to modify the loss functions [14–16], add the pair training [17,18], combine GAN with deep convolution framework [19], adopt multiple GANs and scalable ScoreMix for different applications [20] to improve the image generation quality or speed, which further inspire their applications or new GANs [21–29] with the combination of these techniques for medical images such as bone marrow cell, brain tumors, fundus and cardiac images. However, the good balances of generation quality and speed in these enhanced GANs are mostly ignored. Meanwhile, these intelligence assisted medical imaging methods mostly focus on lung, heart, eyes in magnetic resonance (MRI) images and computer tomography (CT) images, and they cannot be directly transferred to the increasing X-ray breast images for automatic generation. Therefore, we

focus on how to enhance the GAN to achieve rapid automatic generation of high-precision breast X-ray images.

To evaluate experimental results and verify the effectiveness of X-ray images generators, we select universal and applicable metrics as more significant. Peak signal-to-noise ratio (PSNR) [30] is an objective metric for image quality evaluation. Based on the square of the difference between two images, sum and average to obtain the mean square error. Then, the mean square error is used to obtain the corresponding PSNR. The larger the value of the PSNR, the closer the generated image is to the original image. Fréchet inception distance (FID) [31] uses an inception network to extract features from samples, obtaining the feature spaces of the generated and real samples, and modeling the feature spaces. The diversity score of the image is represented by calculating the distance between the corresponding feature spaces. Therefore, considering the difficulties in obtaining medical images and the need for a large number of datasets when combined with artificial intelligence algorithms in the field of medical images, we propose an enhanced GAN, which can accurately and quickly capture multiple tumor information of different sizes and shapes in breast images. The main contributions include:

- We design a fast and accurate framework dm-GAN for automatic breast image generation. The proposed novel dm-GAN model takes advantages of a multi-latent code inversion generator and a multi-discriminator to improve the speed and accuracy of breast image automatic generation.
- We implement the breast image generator dm-GAN on the open-source Digital Database for Screening Mammography (DDSM) dataset to verify its effectiveness.
- We prove that our proposed dm-GAN can automatically generate breast images with higher accuracy (up to 1.84 dB higher PSNR and 5.61% lower FID) and 1.84x faster via the qualitative and quantitative results comparison.

The paper is organized as follows: Section 1 gives the brief introduction while Section 2 depicts the related work. Section 3 details the proposed method of dm-GAN. Section 4 shows the comprehensive results and analysis. The conclusion is summarized in Section 5.

2. Related work

GANs have dominated in the field of automatic image generation and augmentation in the past few years. Various GANs have been designed to further improve the image quality or generation speed since GAN was first proposed to generate images through adversarial learning between the generator and the discriminator. In order to achieve higher accuracy in generating images, people have attempted to optimize the loss function. Nowozin et al. proposed an F-GAN [12] framework for calculating f-divergence, Mao et al. proposed an LSGAN framework for calculating least squares loss [15], and Martin Arjovsky proposed a WGAN framework with Wasserstein distance between two distributions as a loss function [11]. EID-GAN designs a new penalty function to handle the imbalanced data augmentation [16]. Conditional generation is handled by paired data. In order to achieve control of the GAN generated results, Isola et al. proposed the pix2pix framework [18], but it requires paired training data. Therefore, Zhu et al. imitated the generator and discriminator structure of pix2pix and proposed the cycleGAN architecture [17]. Multiple GANs or combined GAN with existing work are considered for higher speed. Two GANs are used to learn the mapping from the source domain to the target domain and from the target domain to the source domain, which can effectively avoid the data pair limits of the pix2pix framework. The dual GAN structure requires more parameters and thus make the training

unstable. To solve the problem of training stability, Chintala et al. proposed a deep convolution framework DCGAN [19] that combines CNN and GAN to improve the stability of model training. To reduce overfitting issues during the training process, the ScoreMix method effectively increases the diversity of data and reduces the overfitting problem using the convex combinations of the real samples [20]. Based on the excellent performance of a series of GAN variants in the field of image generation, various GANs have also been used in the medical field for automatic generation [13] and segmentation to solve practical problems such as highly uneven distribution of medical data and difficulty in assisting doctors in obtaining better diagnosis.

Breast X-ray images can visually present the internal structure of breast tissue, and determine whether a lump is benign or malignant by observing the shape, size and smoothness of the shaded areas in the image. In order to improve the discrimination ability of suspected tumors, the breast X-ray image enhancement algorithm [17] does not to change the relative brightness between the breast image background and suspected disease areas. By enhancing the brightness of the edge areas and performing routine histogram equalization processing, it ensures that most of the collected grayscale values belong to medium to low grayscale areas, thereby enhancing the contrast between the breast background and suspected disease areas. In order to achieve automatic classification and labeling of suspected disease areas in breast X-ray images, Radford et al. [19] establish a generative adversarial network structure based on weak supervised learning for suspected disease areas in pathological images, and screens pathological images with labeled data as a training set. Then, by training the discriminant network structure, relevant features that can distinguish between real data and generated data are extracted to achieve classification and labeling.

For the automatic generation of breast X-ray images, Madani et al. applied the DCGAN architecture with relatively stable training to chest X-ray image synthesis [24], mitigating the problem of pattern collapse during GAN training and improving the quality of image generation. GAN and related variants also contribute to the automatic generation of medical images in other formats such as CT images and MR images. Nie et al. used a novel GAN framework to synthesize CT medical images from MR medical images through the method of context awareness [21]. Wolterink et al. [22] applied the simple pix2pix framework into CT image noise reduction. Jiang et al. [23] used the cycleGAN composed of two original GANs to generate lung CT images and MR images. Hu et al. [25] used Wasserstein distance as the loss function of the WGAN framework for cell-level medical image generation and achieved good results in the task of bone marrow cell classification. AsynDGAN modified the discriminator structure and used the multi-discriminator framework to generate brain tumor images for accurate brain tumor segmentation [26]. 3D brain magnetic resonance imaging is augmented by deep convolutional refined auto-encoding alpha GAN to generate more realistic samples [27]. A novel framework for generating an ICGA (Indocyanine Green Angiography) from a fundus image using GAN is proposed to examine eye diseases [28]. We also finish an adversarial semi-supervised learning framework for cardiac image segmentation [29].

However, these related works on GANs are relatively mature for the generation of MR medical images and CT medical images, but there are less works on the generation of X-ray images, mostly focusing on the segmentation and classification of breast pathology images. Inspired by these efforts to improve the accuracy or speed of image enhancement, we have designed and implemented a fast, accurate and automatic breast X-ray image generation method, dm-GAN. Based on the traditional GAN network structure, the design generator adopts a multi latent code inverse mapping structure, learning the target image inverse mapping to multiple hidden code distributions in the latent space, and

transmitting the effective information provided by the hidden code during the generation process. At the same time, a multi-discriminator structure is designed to effectively alleviate the pattern collapse problem of multiple samples mapping to the same distribution caused by a single discriminant loss, significantly improving the quality of generated images. In addition, asynchronous data parallelism is adopted during the training process to train multiple discriminators in dm-GAN, which accelerate the generation procedure.

3. Proposed method dm-GAN

3.1. dm-GAN framework

The proposed dm-GAN is a distributed multi-latent code inverse mapping generation framework, which can generate high-precision breast X-ray images quickly and automatically as Figure 1 shows.
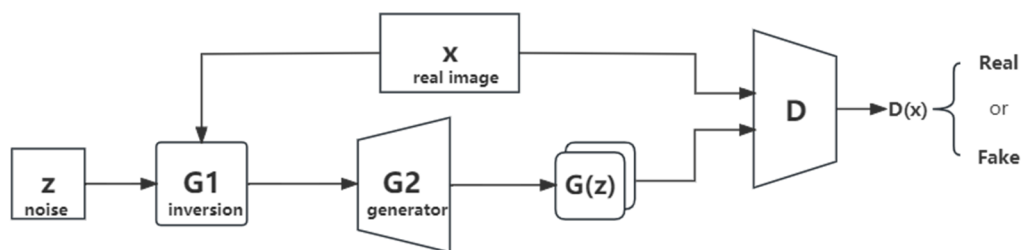

Figure 1. The overall framework of dm-GAN. G1 represents the inverse mapping process, input real image x, noise z and use of G1 network for inversion. G2 learns the image data distribution corresponding to the latent code to generate breast images. Multiple discriminators include {D1, D2,...Dn} while the discriminator learns the different information between the real images and synthetic images from G2.

Based on GAN, dm-GAN adds a multi-latent code inversion method to the generative network so that the generated multiple latent codes can carry more breast image information. Through G2 fitting the real image data distribution corresponding to the multi-latent code in G1, dm-GAN can effectively capture feature details in the breast image. Unlike the discriminator of traditional GAN using single-layer features to express image information, the proposed dm-GAN model adopts a multi-discriminator structure including {D1, D2,…Dn}. Each discriminator first accesses its own data and judges the authenticity of the input image. Then the average loss of multiple discriminators is fed back to the generator, which can effectively prevent the generated images from tending to the same distribution.

3.2. Multi-latent code inversion-based generator

The generator in dm-GAN includes G1 inversion and G2 generator, as is shown in Figure 2. G1 adopts a multi-latent code inverse mapping method. The multi-latent code inverse was proposed in mGAN prior model [32], which uses a pre-trained GAN model to inversely map a given image back to the latent space for multiple latent codes. The previous inverse mapping methods focus on

optimizing the latent code from the perspective of gradient descent [33–35], and learning other encoders to map from the image space back to the latent space [36]. Unlike the single-latent code, multi-latent codes can inversely map more complete feature information into the image. Therefore, G1 adopts the multi-latent code inversion, and the combined latent code after inversion of each picture is the input of the sub-generator G2.
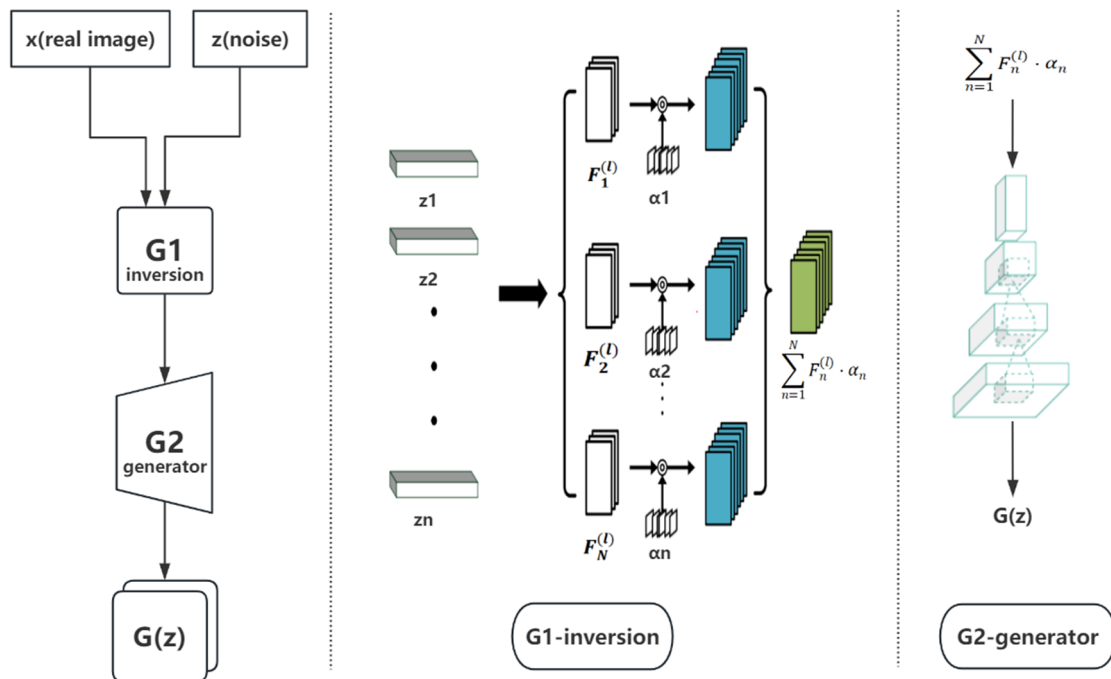
Figure 2. The generators in dm-GAN. The left is the overall generator structure of dm-GAN, while the middle is the G1 inversion, and the right one is the G2 generator. G1 inversion builds the intermediate feature space F (), set the index l for performing feature synthesis, and introduce adaptive channel importance coefficients $\{\partial_n\}_{n=1}^N$ for input variables $\{z_n\}_{n=1}^N$ as weighted and finally output $\sum_{n=1}^N F_n^l \cdot \alpha_n$. G2 generator can further use the output of G1 as the input of G2, construct two convolutional layers with a step size of 2, then increase the number of residual blocks in the middle layer, and finally generate the image $G(z)$ through three layers of convolution (including two layers of transposed convolution).

The G2 generator adopts the deep residual network [36]. First, the G2 generator constructs two convolutional layers with a step length of 2, then increases the number of residual blocks in the middle layer, and finally passes through three convolution layers (including two transposed convolution layers) to generate images. In a word, the proposed dm-GAN model can learn the joint distribution from different datasets belonging to different medical entities, and achieve high-precision breast image generation through the sub-generator G1 and the sub-generator G2.

3.3. Discriminator architecture
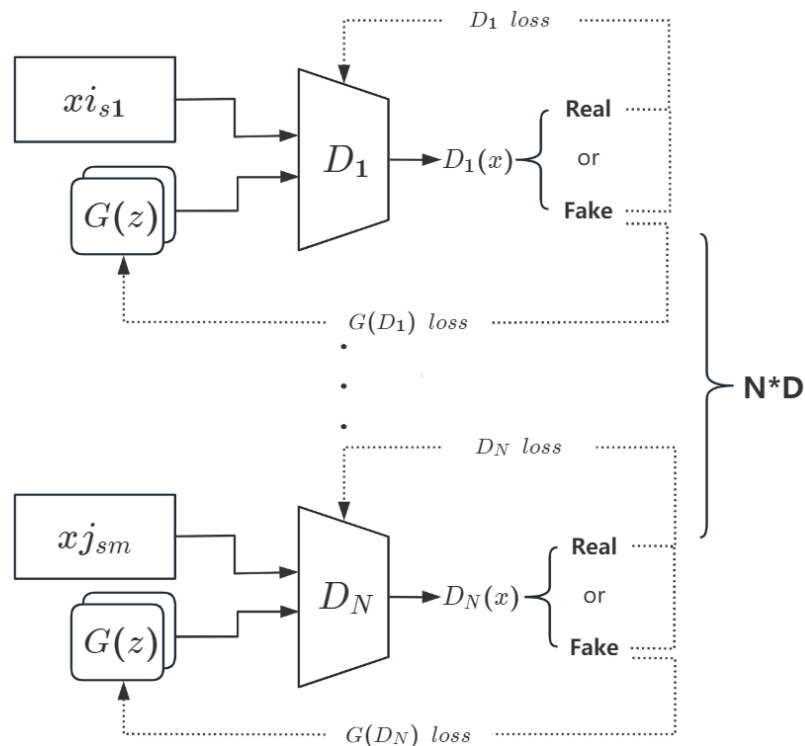
Figure 3. The discriminator structure of dm-GAN. xi_{s1} represents the real image of x_i in the s_1 subset, xj_{sm} represents the real image of x_j in the s_m subset, G(z) represents the false data generated by the generator after inverse learning x and D_1/D_N represents one of the N discriminators, which determines x and G(z) and generates loss values G_loss, D_loss.

To accelerate the generation speed with guaranteed accuracy, multiple discriminators are used in the proposed dm-GAN as shown in Figure 3. The discriminator in GANs could characterize the differences between the generated image and the real image, which affects the accuracy of identifying true and false samples [37]. The proposed dm-GAN uses multiple discriminators, which is inspired by the idea of multiple discriminators in AsynDGAN [26]. The multiple discriminators in the dm-GAN network are composed of multiple independent PatchGAN discriminators. Each discriminator includes multiple convolutional layers. The discriminator judges the generated breast images and the real images by learning the feature distribution in the sub-dataset, and outputs an N-dimensional matrix. Each value in the matrix corresponds to a fixed-size block in the original image [38]. It is assumed that the pixels of each small block are independent [39] in the Markov random field. The quantized output of the true and false values of small blocks helps to capture the high-frequency details in the breast image, and can process images with any size.

The dm-GAN discriminator adopts asynchronous training and makes good use of the data parallel mode for training. Each discriminator accesses the real data stored in its own node and the data generated by the generator and provides a sub-dataset as conditional input. The data generated by the generator is determined by multiple discriminators to determine the probability of belonging to the associated real sub-dataset.

3.4. Objective function of dm-GAN

In the proposed dm-GAN model, G1 uses a pre-trained model to reverse map the target image to the latent space, and generates multiple latent codes that can fit the data distribution of different regions of the target image. G1 optimizes the adaptive channel importance and multi-latent code by calculating the reconstruction loss between the generated image and the original image. The multi-latent code generated by G1 is used as the input of G2 to build the mapping relationship (G ()) between the multi-latent code with real data distribution to generate images. Then the real image x, the generated image G (x) and the real data y in each subset are used as the input of the discriminator. Multiple independent discriminators learn the characteristics of the real data y in each subset to judge the probability that the input image belongs to the subset.

According to the analysis of the input and output of the generator as well as the discriminator in dm-GAN, it is concluded that the real data x in any subset obeys a mixed distribution as follows:

$$s(x) = \sum_{j \in [N]} \pi_j s_j (x) \tag{1}$$

The discriminator D_j corresponding to each sub-dataset receives the data generated by the previous sub-set $s_j(x)$ and the real data in the sub-dataset $s_j(x)$. Therefore, D_j loss is only related to $s_j(x)$, and we return the losses generated by N discriminators and feedback to the generator in Eq (2):

$$\min_{G} \max_{D_1 : D_N} V(D_{1:N}, G) = \sum_{j \in [N]} \pi_j \left\{ E_{x \sim s_j(x)} E_{y \sim p_{data}(y|x)} \left[log \, D_j \, (y|x) \right] \right\} + E_{\hat{y} \sim p\hat{y}(\hat{y}|x)} \left[log \, (1 - \right.$$

$$\left. D_j(\hat{y}|x)) \right] \tag{2}$$

The objective function can be understood as a problem of maximizing and minimizing the optimal solution. The real data x obeys the mixed distribution $s_j(x)$, the conditional input y means that the real data P_{data} is divided into N sub-datasets, $D_{1:N}$ includes $D_1, D_2...D_N,$ all the N independent discriminators, where each sub-discriminator D_j performs data on the data generated by the generator and the real data. The discriminator produces a discriminant loss $logD_j(y|x)$, while the discriminant loss propagates back to the generator to produce a generator loss $log(1-D_j(y|x))$. Nash balance is achieved by optimizing the generator and discriminator losses.

3.5. *dm-GAN algorithm*

The proposed dm-GAN adopts an alternate training method to continuously optimize the parameters of the generator and discriminator models, as Algorithm 1 depicts, so that the final model can generate high-precision breast images very well.

As is shown in Algorithm 1 of our proposed dm-GAN, the line 1 to line 4 give the input of the generator and initialize the parameters. By extracting m sample image datasets X as the input of the generator and initializing N latent codes and weight factors, G1 completes the inversion of m sample images to obtain the corresponding combined latent code, and uses them as the input of the G2 network.

Algorithm 1 Training of dm-GAN

Input: Real image of breast lesions with mass, calcification and so on.

Output: Generated high-precision breast images

---

1) Extract m samples X=$\{\mathbf{x_1^j}, \ldots, \mathbf{x_m^j}\}$ from the mixed distribution $s_j(x)$ as the input of the generator;

2) Initialize N latent codes $\mathbf{z_1, z_2, \ldots, z_N}$, where different latent code captures the features $\mathbf{F_1^l, F_2^l, \ldots, F_N^l}$ of the middle layer of m sample images and initialize the weight factors $\mathbf{\alpha_1, \alpha_2, \ldots, \alpha_N}$;

3) $\sum_{n=1}^{N}\{F_n^{(l)}\}_{i,j,c} \times \{\alpha_n\}_c$; // **Adaptive weighted fusion of middle layer features to maintain semantic consistency**

4) $\sum_{n=1}^{N} F_n^l \odot \alpha_n$; // **Use the obtained multiple sets of hidden codes as input to G2**

5) Set up N discriminators and start training the dm-GAN network model;

6) for epoch in numbers:

7) for k in epoch_D:

8) for D_j in D[N]:

9) m samples X=$\{\mathbf{x_1^j}, \ldots, \mathbf{x_m^j}\}$ corresponding latent code as the input of G2

10) Select m real data Y=$\{\mathbf{y_1^j}, \ldots, \mathbf{y_m^j}\}$ including X as the input of the D_j

11) G generates m sample data as the input of the discriminator D_j

12) Use stochastic gradient ascent to update each discriminator D_j:

$$\nabla_{\theta_{D_j}} \frac{1}{m} \sum_{i=1}^{m} [log D_j(y_i^j) + log(1 - D_j(G(y_i^{'j})))]$$

13) end for D_j; // **End training discriminator Dj**

14) end for D; // **End training discriminator D**

15) for D_j in D[N]:

16) Backpropagate D_loss generated by D_j to the generator G

17) end for D_j; // **End the loop for backpropagation of each discrimination loss to G**

18) Use stochastic gradient descent to update the discriminator G:

$$\nabla_{\theta_G} \frac{1}{Nm} \sum_{j=1}^{N} \sum_{i=1}^{m} log(1 - D_j(G(y_i^{'j}))$$

19) end for nums; // **End the training and iterate a total of nums epochs**

Then, line 5 to line 18 aims at training dm-GAN. For each epoch to train the discriminator, the fixed generator does not perform back propagation. A real image dataset Y including X is established for each discriminator, where Y includes m real images. The real images Y and m images are generated by the generator and are the input to each discriminator. The discriminator makes the decision of the generated image and the real image based on the conditions and feeds back the discriminatory loss to each discriminator. And then it updates the parameters of each discriminator through random gradient rise. Once the first round of training is completed for all discriminators, then the average of the multiple

discriminative losses and their back propagation to each node of the generator is calculated. The parameters of the generator are updated through random gradient descent.

Similarly, the generator and the discriminator continue to update the parameters and optimize the network until all epochs finish in line 19. This section provides a detailed description of dm-GAN algorithm from five parts: the framework, generator structure, discriminator structure, loss function and pseudo code. Among them, the generator uses a multi hidden code inverse mapping structure to more effectively capture the feature details of critical regions in breast images. The multi discriminator structure can avoid the generation of images tending to the same distribution due to single discrimination loss. At the same time, it can protect the privacy of medical data well. In summary, these techniques further demonstrate the feasibility of dm-GAN for fast, accurate and automatic generation of breast X-ray images.

4. Results and analysis

4.1. Experiments configuration

The open-source dataset DDSM [40] is used in these experiments. DDSM includes breast molybdenum target image information of approximately 2500 women from 1988 to 1999, including 10,480 images [41]. These images cover a large number of diverse cases. Therefore, DDSM dataset is often used for breast image research. The cases involved in the dataset are mainly divided into four groups: Normal, Benign, Benign without Callback and Cancer, which represent images of normal, benign, or malignant lesions. This paper mainly focuses on the generation of breast tumor images, so that we select 210 breast lesion images from the Cancer category molybdenum target images and convert them into PNG image format. These images displaying features of lumps or calcifications are selected as the initial input dataset.

Then, the initial dataset is preprocessed as Figure 4 shows. First is to remove redundant background from the original image, and the second step is to randomly crop the image multiple times while keeping the breast intact. The size is set to 256×256, and the contrast-restricted adaptive histogram equalization method is used to perform noise reduction to obtain image enhancement data. The dataset is divided into two groups: training dataset (170 cases) and test dataset (40 cases). And then the training set is classified into 10 subsets on average according to the size of the breasts, and is processed as a subset of the 10 discriminators.

(a)   (b)   (c)   (d)
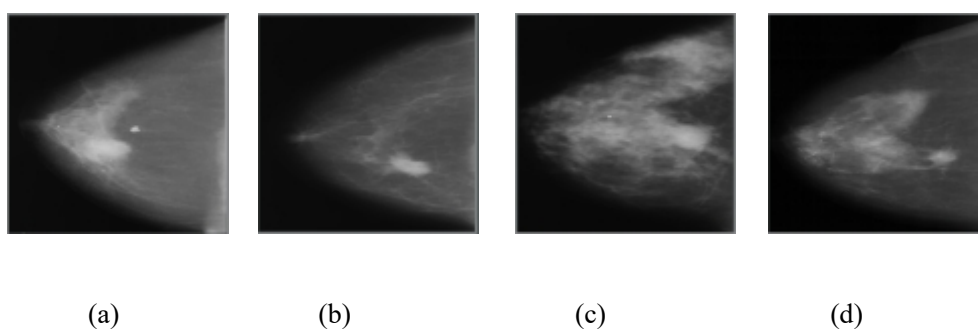
Figure 4. Preprocessing of four different breast lesion images via removing redundant background and cropping the images multiple time.

The generators and discriminators are configured in the proposed dm-GAN. G1 adopts mGANprior_bridge to pretrain the model and refers to network structure Pix2pix [18] to build our dm-GAN. dm-GAN includes multiple convolutional layers and outputs an N-dimensional matrix. Each value in the matrix represents a receptive field in the original image, which corresponds to a fixed size block in the original image. Assuming that each small pixel is independent, Markov random fields are used to quantify the true and false values of different small blocks in the image, which can effectively capture some high-frequency and detailed features in breast images. Compared to the discriminator of a full image input, the size of the tumor is reduced in dimensionality and requires fewer parameters, which can process images with variable size. Multiple discriminators are designed to generate adversarial network structures. the training dataset is divided into 10 subsets, which is the number of discriminators set to 10. The real image is input to the training set, the number of latent codes is set to 20 and the number of medial feature layers to 8. G2 is an encoder-decoder network, including two convolutions for image downsampling, two transposed convolutions and nine residual blocks. The generator will learn a joint distribution from different data sets belonging to different discriminators to generate a 256×256 breast image. The discriminator structure is the same as PatchGAN proposed in the Pix2pix model, and the block size is fixed at 70×70. In the experiment, the Adam optimizer uses a small batch stochastic gradient descent algorithm to optimize the loss function, the momentum parameter is $\beta_1 = 0.5$, $\beta_2 = 0.999$, and the learning rate is 0.0002. The discriminator and generator of dm-GAN are trained by confronting each other. Each discriminator accesses the data stored in each node to distinguish between the generated data and the real data. The generator is a joint learning from different sub-datasets. Thus, the loss value of each discriminator is updated and passed to the generator for further calculating the loss to achieve the objective function optimization.

Once the dm-GAN framework and related parameters are decided, the model is trained and generates breast images. To verify the effectiveness of our proposed dm-GAN, the latest generation methods are chosen as the baselines: CycleGAN[17], DCGAN[19], mGANprior serials[24] (including four mGANprior variants: mGANprior_person, mGANprior_cat, mGANprior_bird and mGANprior_bridge for different scenarios).

### 4.2. Qualitative results and analysis

dm-GAN combines the multi-latent code inversion method with the traditional GAN generator structure, and adopts the dropout method to avoid over-fitting, which helps to improve the accuracy and diversity of generated breast images. Therefore, the qualitative analysis of the diversity and accuracy of the experimental results is listed to verify the effectiveness of proposed dm-GAN over the state-of-the-art in Figure 5.

The outline of the breast lesion image generated by the dm-GAN model in Figure 5(h) is relatively clear and recognizable. More importantly, the shape, size of the mass and area of the calcification image are nearly same as the original images in Figure 5(a).

As for the breast lesion images generated by CycleGAN, there are no related areas, such as the contour or size and shape of the mass, the edge of the calcified area is blurred and the calcified tissue is almost invisible in Figure 5(b). In the breast lesion image generated by DCGAN, the basic attributes such as the contour and size of the tumor can be distinguished and the calcification area is relatively complete in Figure 5(c). However, the shape of the tumor and the generation of calcified tissue are not clear. Therefore, dm-GAN generates better breast images than CycleGAN and DCGAN.
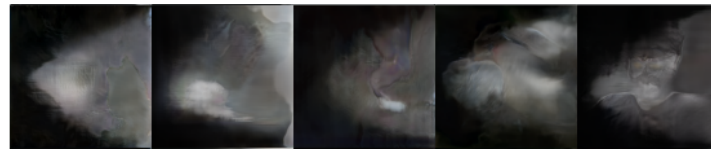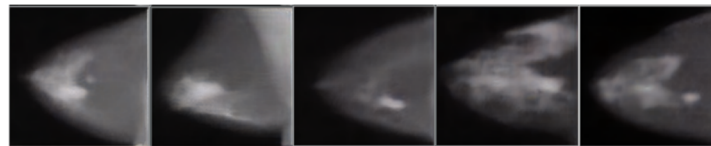
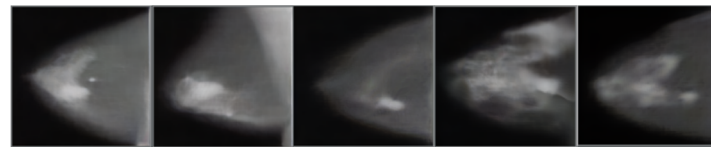(a)   Five original breast images with different shape, size, masses
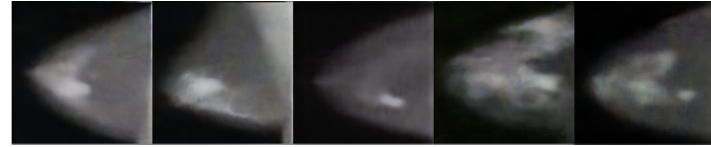
(b)   Five images generated by CycleGAN
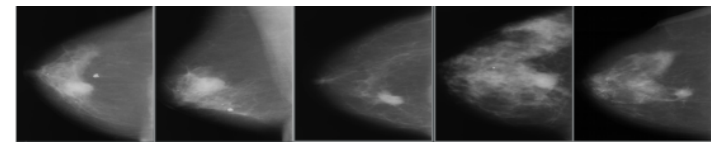
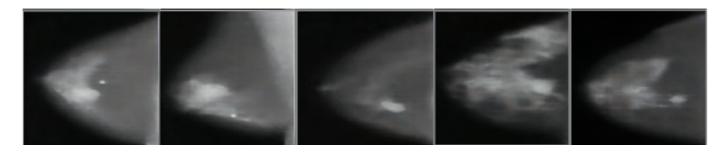(c)   Five images generated by DCGAN

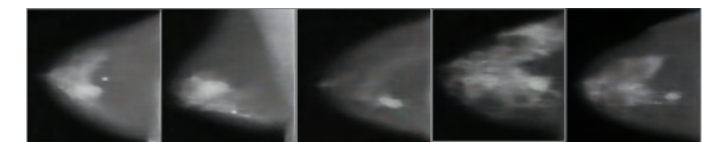(d)   Five images generated by proposed mGANprior_person

(e)   Five images generated by proposed mGANprior_cat

(f)   Five images generated by proposed mGANprior_bird

(g)   Five images generated by proposed mGANprior_bridge

(h)   Five images generated by proposed dm-GAN

Figure 5. Comparison of dm-GAN generation results with existing work.

The serial models of mGANprior are reproduced to generate breast images in Figure 5(d)–(g). The shape, size, contour, calcified area and calcified tissue of the tumor from mGANprior are better than DCGAN and CycleGAN. As for mass images, the mGANprior_bridge model generates the best image quality. The mGANprior_bridge model easily captures the line features in the image, so that the latent code representing the information of the contour of the mass is larger, thus the output image seems to be the same in size and the contour is clearer. Considering the calcified images, mGANprior_cat and mGANprior_bird are slightly better than the mGANprior_bridge. Because the calcified part of the image is larger in area and overlaps with the breast tissue, mGANprior_cat and mGANprior_bird, which are used to capture hair details, can also get good results. It is noted that the different mGAN models have their own advantages in different scenarios. Compared with the proposed dm-GAN, mGANprior_bridge, mGANprior_cat and mGANprior_bird models generate clear size and shape of the tumor in the tumor image, as well as the size and shape of the calcification area.

Through the comprehensive comparison, it is found that the diversity of breast images generated by dm-GAN model performs better. The detailed quantitative results are described in Subsection 4.3.

4.3. Quantitative results and analysis

PSNR and FID are used to evaluate various GANs. The higher PSNR, the higher image quality. To compute the value of PSNR in Eq (3), it is necessary to prepare mean square error (MSE) in Eq (4) [42].

$$PSNR = 10 \times log_{10}(\frac{MAX_I^2}{MSE}) \tag{3}$$

$$MSE = \frac{1}{mn} \times \sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \tag{4}$$

I is the original $m \times n$ image and K is the generated image. In this paper, m and n are both 256. MSE is the mean square error. MAX_I^2 is the maximum pixel value and determined by the bit-width used to express pixels. For example, if 8 bits are used to express pixels, its maximum value is 255 (computed by 2^8-1).

Instead, the smaller the FID value is, the closer the two Gaussian distributions are, and the sharpness of the generated image is higher and the diversity is richer. It is calculated by Eq (5) [31].

$$FID = ||u_r - u_g||^2 + T_r(\textstyle\sum_r + \sum_g - 2 \times (\sum_r\sum_g)^{1/2}) \tag{5}$$

where u_r denotes the features mean of original image and u_g is that of generated image. $\sum_r$ and $\sum_g$ represent the covariance matrix of the original image and generated image, respectively. FID can reflect the mean and variance of the generated image and the target image, and can be calculated by the distance between the two images corresponding to the Gaussian distribution.

According to the qualitative analysis of various GANs in the shape, size, contour, calcified area of generated images and calcified tissue properties of the tumor, dm-GAN and mGANprior models are better than CycleGAN and DCGAN. Therefore, the similarity between the generated image and the original image is analyzed by calculating the PSNR value of the dm-GAN generated image and the PSNR value of the image generated by the different mGANprior models in Table 1. To verify the effectiveness of dm-GAN from the perspective of PSNR, five images are selected and their PSNR values are averaged to compare with various mGANprior models. The PSNR value of the proposed

dm-GAN model achieves 1.31 to 1.84 dB higher than mGANprior models, so that it can generate better quality breast lesion images.

Table 1. PSNR comparison between mGANprior models and dm-GAN.

Models	img1	img2	img3	img4	img5	average
mGANprior_person [24]	36.60	36.00	37.78	33.58	36.42	36.08
mGANprior_cat [24]	38.19	36.69	37.57	35.73	35.69	36.77
mGANprior_bird [24]	37.11	34.82	36.93	32.90	37.08	35.77
mGANprior_bridge [24]	38.24	37.02	38.50	35.70	37.03	37.30
Proposed dm-GAN	**39.53**	**38.97**	**39.79**	**36.51**	**38.24**	**38.61**

The FID score of the image generated by dm-GAN is the lowest at 28.43, up to a 57.49% decrease over CycleGAN in Table 2. Similar to the PSNR results comparison, dm-GAN performs better and has 5.61% lower FID than mGANprior_bridge. The results indicate that the multi-discriminator structure captures the characteristics of the real data in each node and then judges the input image well. The loss values of multiple independent discriminators are averaged and passed to the generator, and the generator learns the mean difference between the image generated after the last iteration and the original image, and the parameters are adjusted and optimized, which can achieve high-precision and diverse breast image generation. dm-GAN alleviates the over-fitting problem by randomly ignoring some neurons during forward propagation in G2 and avoiding all image features at once. If a multi-discriminator structure is set up, over-fitting caused by training small datasets can be effectively addressed by an asynchronous training method.

Table 2. FID comparison between various models.

Models	FID	Decrease%
mGANprior_bridge [24]	31.56	51.08%
DCGAN [19]	35.89	44.37%
CycleGAN [17]	64.52	0
Our proposed dm-GAN	**28.43**	**57.49%**

At the same time, since each discriminator in dm-GAN can only access the data in its own node, it can effectively protect the privacy of medical data.

The GAN-based deep learning methods are used to generate breast images, and their image generation time can effectively explain the generation speed. Here the generation time comparison of different models is listed in Table 3 to illustrate the performance of the dm-GAN model. It can be seen from the table that the image generation time of the mGANprior_bridge model is the longest, up to 110 hours for 100 epochs, which takes about 1.1 hours for each epoch. DCGAN and CycleGAN are different from the traditional GAN structures, and the time to complete the generation of 210 images for 100 epoches is about 1/4 of the mGANprior model, only around 30 hours. The proposed dm-GAN model combines the multi-latent code inverse mapping method and the multi-discriminator structure to tradeoff the accuracy and speed well. It can achieve a 1.38x speedup over the mGANprior model. Therefore, it can be concluded that the dm-GAN model helps to improve the accuracy and diversity of mammary X-ray images, and improves the generation efficiency as well.

Table 3. Generation time comparison of models.

Models	Time (hours)	Speedup
mGANprior_bridge [24]	110	1
DCGAN [19]	25	4.4
CycleGAN [17]	30	3.67
Our proposed dm-GAN	**80**	**1.38**

It is noted that the proposed dm-GAN is very useful not only for breast X-ray images, but also suitable for other kinds of images, such as optical coherence tomography (OCT) [43,44]. Therefore, we plan to further apply this novel dm-GAN to generate OCT images and assist retinitis pigmentosa diagnosis generation.

5. Conclusions

We propose a fast, accurate and automatic beast image generator, dm-GAN, using distributed multi-latent code inverse mapping. On one hand, the dm-GAN generator adopts the multi-latent code inverse mapping method to improve the learning ability of fitting the real data distribution to generate high-precision breast images. On the other hand, the dm-GAN discriminator includes ten distributed discriminators, which independently access the data in their respective nodes and generate multiple discriminative losses, so that it can learn more accurate breast feature information and quickly improve the accuracy of the generated image. The comprehensive experiments demonstrate that dm-GAN can achieve up to 1.84 dB higher PSNR, 5.61% higher FID, and a 1.38x speedup compared to the latest mGANprior. However, our proposed dm-GAN has some limitations such as the loss of information in the breast tissue during the inverse mapping process, while the large number of model parameters requires extra training time and cost. In future work, we will continue to optimize the breast image generation method dm-GAN and combine it with effective segmentation and classification for real intelligent diagnosis in hospitals.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Funding

Conflict of interest

The authors declare there is no conflict of interest.

References

1. S. P. Zuckerman, B. L. Sprague, D. L. Weaver, S. Herschorn, E. Conant, Multicenter evaluation of breast cancer screening with digital breast tomosynthesis in combination with synthetic versus digital mammography, *Radiology*, **297** (2020), 545–553.

2. R. Shi, Q. Yao, L. Wu, J. Xu, Breast lesions: diagnosis using diffusion weighted imaging at 1.5 T and 3.0T—systematic review and meta-analysis, *Clin. Breast Cancer*, **18** (2018), 305–320. https://doi.org/10.1016/j.clbc.2017.06.011

3. E. A. Rafferty, J. M. Park, L. E. Philpotts, S. Poplack, J. Sumkin, E. Haipern, et al., Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: Results of a multicenter, multireader trial, *Radiology*, **266** (2013). https://doi.org/10.1148/radiol.12120674

4. M. J. Li, Y. C. Yin, J. Wang, Y. F. Jiang, Green tea compounds in breast cancer prevention and treatment, *World J. Clin. Oncol.*, **5** (2014), 520–528. http://doi.org/10.5306/wjco.v5.i3.520

5. R. Shu, Principles and clinical applications of computer-aided diagnosis (CAD) (in Chinese), *Chin. J. CT MRI*, **2** (2004). https://doi.org/10.3969/j.issn.1672-5131.2004.02.016

6. D. Ribli, A. Horváth, Z. Unger, P. Pollner, I. Csabai, Detecting and classifying lesions in mammograms with deep learning, *Sci. Rep.*, **8** (2018), 4165. https://doi.org/10.1038/s41598-018-22437-z

7. J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, (2015), 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965

8. J. Arevalo, F. A. González, R. Ramos-Pollán, J. L. Oliveira, M. Lopez, Representation learning for mammography mass lesion classification with convolutional neural networks, *Comput. Methods Programs Biomed.*, **127** (2016), 248–257. https://doi.org/10.1016/j.cmpb.2015.12.014

9. M. Zhang, J. Huang, X. Xie, C. D'Arcy J. Holman, Dietary intakes of mushrooms and green tea combine to reduce the risk of breast cancer in Chinese women, *Int. J. Cancer*, **124** (2009), 1404–1408. https://doi.org/10.1002/ijc.24047

10. Ian J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., Generative Adversarial Nets, in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, (2014), 2672–2680. https://dl.acm.org/doi/10.5555/2969033.2969125

11. M. Arjovsky, S. Chintala, L. Bottou, Wasserstein Generative Adversarial Networks, in *Proceedings of the 34th International Conference on Machine Learning*, (2017), 214–223. https://dl.acm.org/doi/abs/10.5555/3305381.3305404

12. M. Mirza, S. Osindero, Conditional Generative Adversarial nets, *arXiv preprint*, (2014), arXiv:1411.1784. https://doi.org/10.48550/arXiv.1411.1784

13. X. Yi, E. Walia, P. Babyn, Generative Adversarial Network in medical imaging: A review, *Med. Image Anal.*, **58** (2019), 101552. https://doi.org/10.1016/j.media.2019.101552

14. S. Nowozin, B. Cseke, R. Tomioka, F-GAN: training generative neural samplers using variational divergence minimization. in *Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS'16)*, (2016), 271–279. https://dl.acm.org/doi/10.5555/3157096.3157127

15. X. Mao, Q. Li, H. Xie, R. Y. K. Lau, Z. Wang, S. P. Smolley, Least squares Generative Adversarial Networks, in 2017 *IEEE International Conference on Computer Vision (ICCV)*, (2017), 2813–2821. https://dl.acm.org/doi/10.1109/ICCV.2017.304

16. W. Li, J. Chen, J. Cao, C. Ma, J. Wang, X. Cui, et al., EID-GAN: Generative Adversarial Nets for extremely imbalanced data augmentation, *IEEE Trans. Ind. Inf.*, **19** (2023), 3208–3218. https://doi.org/10.1109/TII.2022.3182781

17. J. Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using Cycle-Consistent Adversarial Networks, in *IEEE International Conference on Computer Vision (ICCV)*, (2017), 2242–2251. https://doi.org/10.1109/ICCV.2017.244

18. P. Isola, J. Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with Conditional Adversarial Networks, in 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017), 5967–5976, https://doi.org/10.1109/CVPR.2017.632

19. A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with Deep Convolutional Generative Adversarial Networks, *arxiv preprint*, (2015), arXiv:1511.06434. https://doi.org/10.48550/arXiv.1511.06434

20. J. Cao, M. Luo, J. Yu, M. H. Yang, R. He, ScoreMix: A scalable augmentation strategy for training GANs with limited data, *IEEE Trans. Pattern Anal. Mach. Intell.*, **45** (2023), 8920–8935. https://doi.org/10.1109/TPAMI.2022.3231649

21. D. Nie, X. Cao, Y. Gao, L. Wang, D. Shen, Estimating CT image from MRI data using 3D fully convolutional networks, in *Deep Learning and Data Labeling for Medical Applications*, Springer, (2016). https://doi.org/10.1007/978-3-319-46976-8_18

22. J. M. Wolterink, T. Leiner, M. A. Viergever, I. Išgum, Generative Adversarial Networks for noise reduction in low-dose CT, *IEEE Trans. Med. Imaging*, **36** (2017), 2536–2545. https://doi.org/10.1109/TMI.2017.2708987

23. J. Jiang, Y. C. Hu, N. Tyagi, P. Zhang, A. Rimner, G. S. Mageras, et al., Tumor-aware, adversarial domain adaptation from CT to MRI for lung cancer segmentation, in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018*, (2018), 777–785. https://doi.org/10.1007/978-3-030-00934-2_86

24. A. Madani, M. Moradi, A. Karargyris, T. Syeda-Mahmood, Semi-supervised learning with generative adversarial networks for chest X-ray classification with ability of data domain adaptation, in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, (2018), 1038–1042, https://doi.org/10.1109/ISBI.2018.8363749

25. B. Hu, Y. Tang, E. I. C. Chang, Y. Fan, M. Lai, Y. Xu, Unsupervised learning for cell-level visual representation in histopathology images with Generative Adversarial Networks, *IEEE J. Biomed. Health. Inf.*, **23** (2019), 1316–1328. https://doi.org/10.1109/JBHI.2018.2852639

26. Q. Chang, H. Qu, Y. Zhang, M. Sabuncu, C. Chen, T. Zhang, et al., Synthetic learning: Learn from distributed asynchronized discriminator GAN without sharing medical image data, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2020), 13853–13863. https://doi.org/10.1109/CVPR42600.2020.01387

27. A. Segato, V. Corbetta, M. D. Marzo, L. Pozzi, E. De Momi, Data augmentation of 3D brain environment using deep convolutional refined auto-encoding alpha GAN, *IEEE Trans. Med. Rob. Bionics*, **3** (2021), 269–272. https://doi.org/10.1109/TMRB.2020.3045230

28. P. Tanachotnarangkun, S. Marukatat, I. Kumazawa, P. Chanvarasuth, P. Ruamviboonsuk, A. Amornpetchsathaporn, et al., A framework for generating an ICGA from a fundus image using GAN, in *2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, (2022), 1–4, https://doi.org/10.1109/ECTI-CON54298.2022.9795543

29. W. Cheng, J. Jiao, An adversarially consensus model of augmented unlabeled data for cardiac image segmentation (CAU+), *Math. Biosci. Eng.*, **20** (2023), 13521–13541. https://doi.org/10.3934/mbe.2023603

30. D. Pan, L. Jia, A. Zeng, X. Song, Application of generative adversarial networks in medical image processing, *J. Biomed. Eng.*, **35** (2018), 970–976. https://doi.org/10.7507/1001-5515.201803025

31. D. C. Dowson, B. V. Landau, The Fréchet distance between multivariate normal distributions, *J. Multivar. Anal.*, **12** (1982), 450–455. https://doi.org/10.1016/0047-259X(82)90077-X

32. J. Gu, Y. Shen, B. Zhou, Image processing using multi-code GAN prior, in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2020), 3009–3018. https://doi.org/10.1109/CVPR42600.2020.00308

33. Z. Lipton, S. Tripathi, Precise recovery of latent vectors from Generative Adversarial Networks, *arXiv preprint*, (2017), arXiv:1702.04782. https://doi.org/10.48550/arXiv.1702.04782

34. A. Creswell, A. A. Bharath, Inverting the generator of a Generative Adversarial Network, *IEEE Trans. Neural Networks Learn. Syst.*, **30** (2019), 1967–1974. https://doi.org/10.1109/TNNLS.2018.2875194

35. F. Ma, U. Ayaz, S. Karaman, Invertibility of Convolutional Generative Networks from partial measurements, in *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18)*, (2018), 9651–9660. https://dl.acm.org/doi/10.5555/3327546.3327632

36. G. Perarnau, J. van de Weijer, B. Raducanu, J. M. Álvarez, Invertible conditional GANs for image editing, *arXiv preprint*, (2016), arXiv:1611.06355. https://doi.org/10.48550/arXiv.1611.06355

37. D. Bau, H. Strobelt, W. Peebles, J. Wulff, B. Zhou, J. Zhu, et al., Semantic photo manipulation with a generative image prior, **38** (2019), 1–11. https://doi.org/10.1145/3306346.3323023

38. D. P. Kingma, P. Dhariwal, Glow: Generative flow with invertible 1x1 convolutions, *arXiv preprint*, (2018), arXiv:1807.03039. https://doi.org/10.48550/arXiv.1807.03039

39. C. Li, M. Wand, Precomputed real-time texture synthesis with Markovian Generative Adversarial Networks, in *Computer Vision–ECCV 2016*, Springer, (2016). https://doi.org/10.1007/978-3-319-46487-9_43

40. M. Heath, K. Bowyer, D. Kopans, The digital database for screening mammography, in *Proceedings of the 5th International Workshop on Digital Mammography*, (2000), 212–218.

41. M. Benndorf, C. Herda, M. Langer, E. Kotter, Provision of the DDSM mammography metadata in an accessible format, *Med. Phys.*, **41** (2014), 051902. https://doi.org/10.1118/1.4870379

42. K. Chen, Q. Qiao, Z. Song, Applications of Generative Adversarial Networks in medical images (in Chinese), *Life Sci. Instrum.*, **Z1** (2018).

43. R. K. Meleppat, P. Zhang, M. J. Ju, S. K. K. Manna, Y. Jian, E. N. Pugh, et al., Directional optical coherence tomography reveals melanin concentration-dependent scattering properties of retinal pigment epithelium, *J. Biomed. Opt.*, **24** (2019). https://doi.org/10.1117/1.JBO.24.6.066011

44. D. Sakai, S. Takagi, K. Totani, M. Yamamoto, M. Matsuzaki, M. Yamanari, et al., Retinal pigment epithelium melanin imaging using polarization-sensitive optical coherence tomography for patients with retinitis pigmentosa, *Sci. Rep.*, **12** (2022). https://doi.org/10.1038/s41598-022-11192-x