Article | Received 30 September 2024; Accepted 18 February 2025; Published 27 February 2025
https://doi.org/10.55092/aimat20250007

EPG-GAN: edge-guided progressive growing generative adversarial network for mammogram synthesis

Shangyu Wang and Shaoping Wang *

School of Automation Science and Electrical Engineering, Beihang University, Beijing, China

* Correspondence author; E-mail: shaopingwang@buaa.edu.cn.

Highlights:

- EPG-GAN enhances texture customization for realistic material properties in mammograms.
- The model's high-resolution synthesis mimics material microstructure details accurately.
- Edge guidance aids in generating structural features vital for material characterization.

Abstract: Mammography is widely used for early detection of breast cancer in clinical diagnosis. The training of new radiologist and deep-learning based algorithms for mammogram analysis usually requires a large amount of data. However, obtaining and labeling large-scale mammogram data are not easy tasks, especially for diseases with low incidence. Realistic mammogram synthesis can alleviate this problem to a great extent. In this paper, we propose an edge-guided progressive growing generative adversarial network (EPG-GAN) to synthesize realistic mammograms with high-resolution and customized texture editing features. By introducing the design of auxiliary edge guidance into the EPG-GAN, the model is empowered to enhance structural details of generated images and avoid synthesizing unnatural images. We superpose the edge sketch onto the object mask and use the composite mask as the network input. In addition, for generating high resolution mammograms, we adopt a progressive training strategy to gradually generate high-resolution images from low-resolution images. To demonstrate the effectiveness of our proposed model, we conduct experiments on publicly available INbreast dataset. Qualitative and quantitative evaluations validate the performance of our model in synthesis high-resolution mammogram. The proposed EPG-GAN achieves 0.983 in structure similarity index measure (SSIM) score, which is significantly higher than other models. The results show that the EPG-GAN can effectively synthesis realistic and high-resolution mammograms, enabling the augmentation of the dataset and laying the foundation for enhancing CAD for breast cancer.

Keywords: mammogram; generative adversarial network; image synthesis; breast cancer; data augmentation; progressive growing; deep learning

1. Introduction

Deep learning (DL) models have been widely and successfully applied to solving various problems in our modern life. However, the performance of most deep-learning based methods are bottlenecked by the quality and quantity of the data. Several works [1–3] have shown that the generalization error of neural networks reduces linearly with the log of the dataset size. Unfortunately, create a high quality and large dataset is time-consuming and expensive, especially in medical.

Mammogram is a high-resolution photograph of woman's breast taken using x-rays, it is clinically used as the standard breast cancer screening exam for the general population and has been shown effective in early detection of breast cancer and in reduction of mortality [4–6]. Conventional 2D mammography involves digitized film-screen mammography (DFM) that used in early decades and full-field digital mammography (FFDM) that rose in recent decades [7]. Standard mammography views are bilateral craniocaudal (CC), extracted from top-down, and mediolateral-oblique (MLO), an oblique view taken under 45° [8]. These two views comprise routine screening mammography. During diagnosis, doctor first manually operate an imaging equipment to produce images required for diagnosis, and then review and analyze the images to find abnormalities [9]. This process relies heavily on doctors' knowledge and experience. It usually takes a long time for novices to acquire operating and diagnostic skills. This is even truer when diagnosing rare diseases, due to the lack of training on real data [10].

In recent years, we have witnessed considerable progress in computational medical image analysis for the detection, diagnosis, and treatment of diseases [11]. Many computer-aided diagnosis (CAD) methods [12–17] base on the mammogram datasets [18–20] has been widely reported with promising diagnostic accuracy in analyzing breast cancer. Compared with medical image interpretation by human experts, automated analysis is more efficient, objective, and does not suffer from inter-observer variations [21]. In the stream of applying machine learning, especially deep learning, to data analysis, large-scale datasets and annotations lie at the heart of its success to accomplish target tasks. For example, the ImageNet database, designed for visual object recognition, contains more than one million annotated images. However, in real clinical applications, usually only a very limited number of images are available due to privacy concerns, and annotations require expert knowledge about the data and task. Although some new datasets [22–24] have large-scale, few researchers take them due to it's hard to be access. Therefore, the lack of easy-access large-scale datasets and annotations remains a major obstacle hindering the successful application of deep learning algorithms to medical images [25–26].

Researchers have been trying to circumvent this obstacle via data augmentation. The most common method is an affine transformation, including translation, rotation, and scaling. This technique simply modifies original images to expand the dataset for model training. Although the sample size can be remarkably increased in this way, only little additional information is introduced into the dataset, due to the small content changes (e.g. rotating an image by an angle) [27–28]. In this regard, there is an urgent need for a new data augmentation method that can enrich the dataset with more variability, so that the model trained on a small dataset can also generalize well on unseen data [29].

Image synthesis is a new and more sophisticated augmentation method. It can be classified into physics-based and learning-based methods. Well-known X-ray simulation packages such as VICTRE [30] can be used to simulate mammograms through Monte Carlo X-ray transport algorithm. The implemented x-ray interaction physics models are based on those used in the PENELOPE code [31], which are

recognized as one of the most accurate in the field of particle transport simulation. They are very time-consuming, especially for generating high-resolution images. Moreover, their performance may be affected by the quality of pre-built models, which are often essential and difficult to construct.

In the past few years, deep learning based synthetic methods have gained more and more interest. Generative adversarial networks (GANs) [32] have made a significant progress in natural image synthesis and data augmentation (especially conditional-GANs (cGANs) [33]). Many researchers have also applied GAN based model in medical image synthesis such as chest X-rays generation [34] and liver lesions synthesis [35]. Nevertheless, most of the researches [36–39] on GAN-based medical image synthesis focus on generating low-resolution medical samples rather than synthesis high-resolution medical image like mammograms for augmentation of the original dataset to alleviate the issues caused by the insufficiency of data. Shen *et al.* [40] proposed an Infilling-GAN to synthesis 256×256 mammogram with contextual information that achieved the Learned Perceptual Image Patch Similarity (LPIPS) score at 0.025. Oyelade *et al.* [41] trained a ROImammo-GAN to synthesis the mass regions based on mammograms that achieved the Structure Similarity Index Measure (SSIM) score at 0.800 and the Peak Signal-to-Noise Ratio (PSNR) score at 27.72. Joseph *et al.* [42] proposed a prior-guided GAN to synthesis 256×256 mammograms that achieved the Fréchet Inception Distance (FID) score at 2.20. Although cGAN is effective and enables the user-controlled image generation, the synthesized images often have low resolution and checkerboard artifacts. To make the structural details of generated images more realistic, auxiliary guidance information, such as the sketch and edge of the background, was introduced [43–44]. However, it is still challenging to synthesize high-resolution images. Due to the more details in high-resolution images, the discriminator can easily recognize the differences between generated and real images, which may lead to the vanishing gradient problem and make the training difficult. Additionally, training such model is memory intensive, which limits using a large batch size to improve training stability.

To address the issue mentioned above and construct a synthesis model with remarkable performance for synthesis high-resolution mammograms, the edge-guided progressive growing idea was introduced in this study. We proposed a novel Edge-guided Progressive Growing Generative Adversarial Network (EPG-GAN) to synthesize high-fidelity and high-resolution mammograms from simple segmentation maps. In order to enhance the structure fidelity and reduce the time that in synthesis process, we first added fine-grained edge sketch to original label maps, which resulted in the composite label maps that can help the generator create images with realistic texture. With the edge-guided, the model can generate mammograms effectively and avoid synthesizing unnatural images in which the edge has blurred areas and distorted structures. To achieve high-resolution and avoid the use of multiple cGANs (hard to train), we inspired from the Progressive Growing Generative Adversarial Networks (PG-GANs) [45] and redesigned a progressive growing structure to gradually generate high-resolution mammograms. Image will be trained firstly at 128×128 resolution. Then the trained 128×128 resolution image will be transferred to next step's input. By using the Fusion Blocks (FB) for transition, the 128×128 resolution image up-sampling to 256×256 resolution smoothly so that the generator and the discriminator growing better stability. Finally, EPG-GAN can successfully learn to synthesize 512×512 resolution images with less training time.

The contributions of the EPG-GAN are as follows:

(1) We propose a novel GAN-based model for generating high-resolution mammograms. To enhance the fidelity of synthesized structure details, we propose to introduce auxiliary sketch guidance into a cGAN. Specifically, we superpose the edge sketch onto the object mask and use the composite mask as the network input. Customized editing of the edge sketch and object mask makes our method quite flexible in generating different mammograms for augmenting data in deep learning models.

(2) Compared with conventional GANs for mammograms generation, progressive growing structure is incorporated in the proposed EPG-GAN to gradually generate high-resolution images from low-resolution images. It can leverage the fusion block to fade the previous layer in the next layer smoothly.

2. Methods

The proposed EPG-GAN is designed for the augmentation of high-resolution mammogram datasets. EPG-GAN aims to generate realistic and high-resolution mammograms efficiently, which could help to improve the performance of the data-driven computer-aided diagnosis methods. As shown in Figure 1, in EPG-GAN, the edge-guide module extracts the edge information of origin mammograms and guided the generator to generate realistic mammograms. For synthesis high-resolution mammograms, the EPG-GAN has also equipped a progressive growing strategy, which make us start the training of EPG-GAN on low resolution conditional input and then progressively increase the input resolution and add learnable layers to the existing network with sharing weights for continuous training.

2.1. Edge-guided

There are two common image-to-image translation tasks in the field of medical image analysis: translation between different imaging modalities and transformation from segmentation maps to medical images. Image synthesis from segmentation maps can generate different images by simply modifying the content in the maps, which is a desirable feature for data augmentation. Since segmentation maps only contain the shape of target structures and lack background details, the transformation from segmentation maps to medical images is generally more difficult than translation between different imaging modalities. As for mammograms, synthesizing from segmentation maps is even harder due to high-resolution mammograms usually contain complex backgrounds. We used the edge information of the background texture as auxiliary sketch guidance to achieve high-fidelity synthesis and customized editing of mammograms.

Specifically, the Canny algorithm [46] was applied to real images to extract binary edge because it is robust against noise. We then updated the original segmentation map of the target object S by superposing the edge sketch E onto it, resulting in the composite label S, which is defined as:

$$S = (1 - M) \otimes E + M \otimes S \quad (1)$$

where M ($M \in \{0,1\}$) denotes the binary map indicating the area for annotated structures. $\otimes$ refers to the operation of element-wise multiplication. Through the above operation, the auxiliary edge sketch E is superposed onto the original mask S without affecting the area of the target objects. With the

additional auxiliary sketch of background provided for GAN to learn, our method can generate images with realistic background texture.

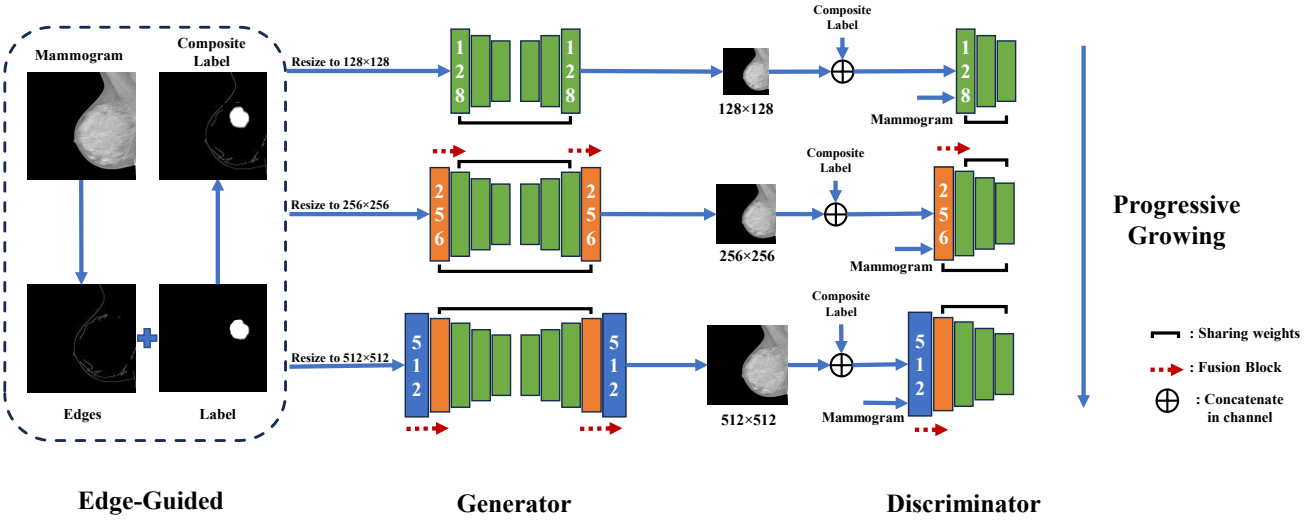

Figure 1. The framework of the proposed Edge-guided Progressive Growing Generative Adversarial Network (EPG-GAN) model. The composite labels contain original annotated structures as well as auxiliary edge sketches. For the generator, the input is the composite labels and output is the generated mammograms. For the discriminator, the input is the mammograms and the corresponding composite labels. In the progressive training scheme, the top backbone structure is adopted as a pre-trained model for low-resolution image (128×128) synthesis, and then fusion blocks are added for synthesizing realistic high-resolution images (256×256).

2.2. Progressive growing and fusion block

Compared to low-resolution image, high resolution images have more fine structures. It is difficult to synthesis high-resolution mammograms due to a high-resolution image amplifies the flaws in structure details and the discriminator can easy to capture. The backbone architecture alone is not capable of extracting enough information for generating high-resolution, realistic images. To tackle the challenges and avoid heavy computation, we propose to adopt the progressive growing training scheme [45] to decompose the task as incremental learning ones. This scheme enables us to use only one generator and one discriminator with fast and smooth learning for high-resolution, realistic synthesis. Specifically, we started from an easier task that synthesizes low resolution images in several warm-up epochs with the backbone structure, and then, the weights of the backbone structure were shared with the generator and discriminator for high-resolution mammograms synthesis.

The entire training process can be divided into four steps. In first step, EPG-GAN is trained with resolution 128×128 edge-guided sketch until convergence, this pre-trained architecture enabled good quality synthesis of low-resolution mammograms. In step 2 and 3, we trained the discriminator and generator for high-resolution image synthesis, sequentially and respectively. In this process, new layers were added to the networks, and we faded them in smoothly with the fusion blocks to avoids sudden shocks to the already well-trained layers. The discriminator was trained earlier than the generator to replenish the gradient information and force the generator to learn to synthesize higher resolution images. Finally, in step 4, we trained the discriminator and generator together for several more epochs

to further enhance performance. By gradually growing the layout of both generator and discriminator for capacity enhancement, the scheme enables EPG-GAN to use only one generator and discriminator with fast and smooth learning for high-resolution, realistic synthesis.

To avoid the sudden shock during training when adding new layers, we adopt the Fusion Blocks for smooth transition in both generator and discriminator. Figure 2 illustrates the structure details of Fusion Blocks. For the purpose of transition, Fusion Block follows the residual design. Facing with the different resolutions between the input and internal layers, Fusion Block applies the added convolution layers and resizes the input in the main stream to match the output resolution. At the same time, Fusion Block introduces a skip connection to directly resize the input and skips it to merge with the main stream through the weight $\alpha \sim (0,1)$. The weight α controls the balance between the main and side branches. $\alpha = 1$ means the output only depends on the main branch, while $\alpha = 0$ indicates that only the side branch determines the output of fusion block. During training, the transition happens as the α increases and the output depends less on original input resolution. The advantages of Fusion Block are that it not only smooths the transition between different resolutions, but also remains the base model structure, making weights sharing possible and hence reducing training time.

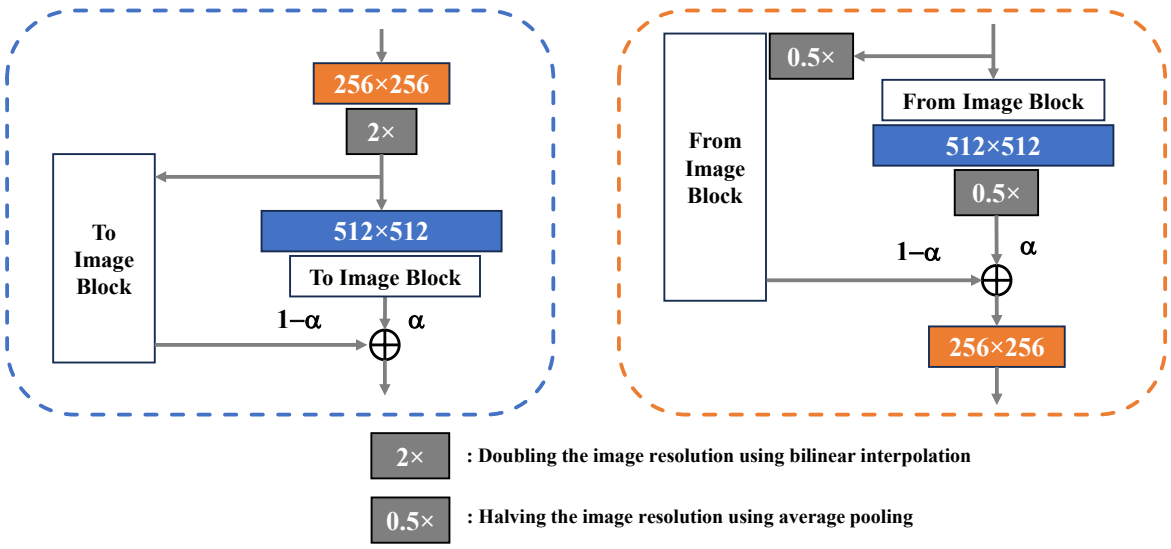

Figure 2. Structure of Fusion Blocks. These two blocks are used in the progressive growing scheme, with α increasing during transition phases. The from image block represents a layer projecting image channels to feature vectors using 1×1 convolution, and the to image block functions the opposite way. The block 512×512 contains two 3×3 convolution layers, and the block 256×256 represents the original structures of backbone adjacent to newly added structures.

Figure 2 shows the two-branch structure in fusion block, there are two types of Fusion Block. One Fusion Block is for down-sampling, which is used in both generator and discriminator. The other Fusion Block is for up-sampling, which is only used in generator. The lightweight side branch helps the pre-trained network adapt easily. The more complex architecture of the main branch has stronger feature extraction capabilities. Combining these advantages, the α is introduced to guide the side branch to gradually switch to the main one. Specifically, it is increased from 0 to α_{max} with a fixed step. It is noted that updating the α in both discriminator and generator simultaneously may decrease the stability and thus we increased α

alternately when training the two modules. Besides, we also noticed that α_{max} had an impact on the performance of the generator. A lower α_{max} may limit feature extraction capabilities, while a larger one tends to cause a sudden increasing loss. In comparison, the loss in the discriminator varies smoothly even with a larger α_{max}. For these reasons, we set the α_{max} to 0.5 and 1.0 in generator and discriminator, respectively.

2.4. Network architectures

We adopt a U-Net-based CNN network as architecture of generator. U-Net [47] architecture is an encoder-decoder network in general, except that skip connections are added between mirrored layers in the down-sampling and up-sampling block. The result of the up-sampling is connected to the feature map of the corresponding layer in the down-sampling path by skip connection. Figure 3 shows the architecture of the generator.

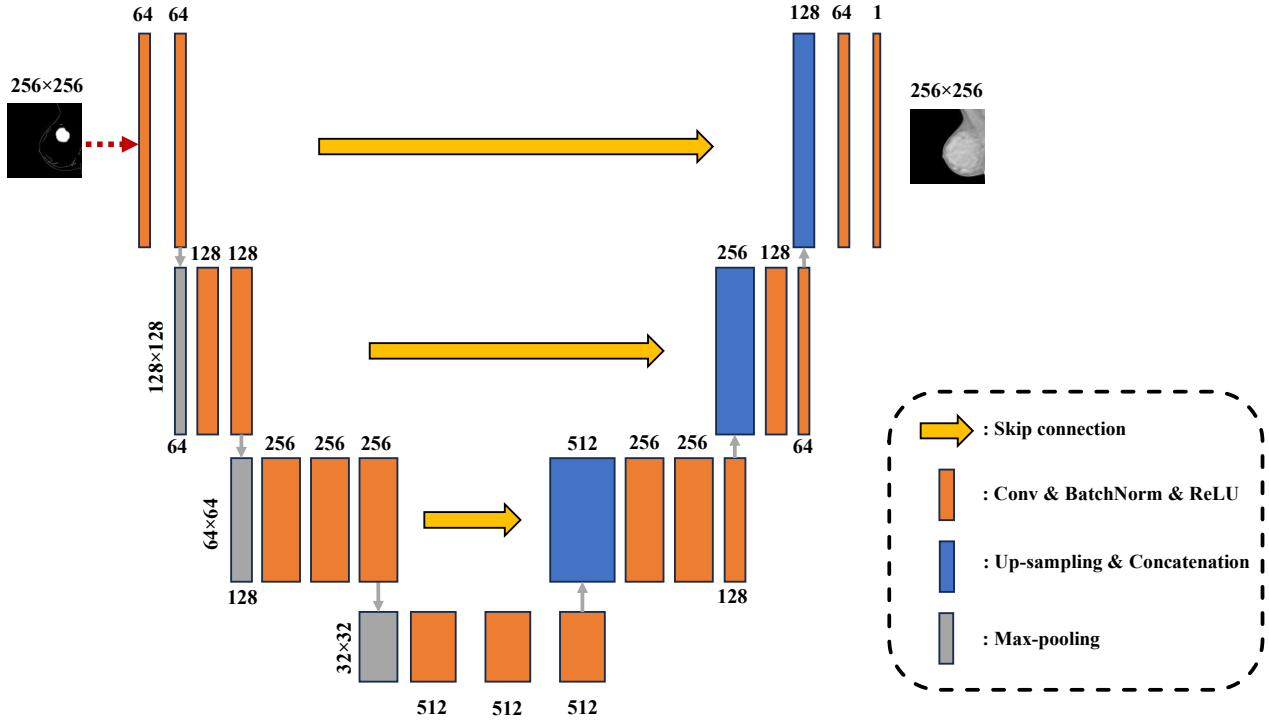

Figure 3. 256×256 resolution step's generator architecture.

The discriminator is based on patch-GAN [48] that considers equisized patches from an image and checks whether the patches are real or fake. As shown in Figure 4, the discriminator takes an image and passes it through a series of convolutional layers, transforming it into a feature vector of size 30×30. The final feature vector is a 2-dimensional vector with values between 0 and 1 representing the probability that each patch is real or fake. If the patch is real, the discriminator is trained to produce an output value close to 1, indicating a high probability that the patch is real. Conversely, if the patch is fake, the discriminator is trained to produce an output value close to 0, indicating a low probability that the patch is real. The convolutional layers in the discriminator typically use small filters and are often followed by batch-normalization and activation functions such as LeakyReLU.

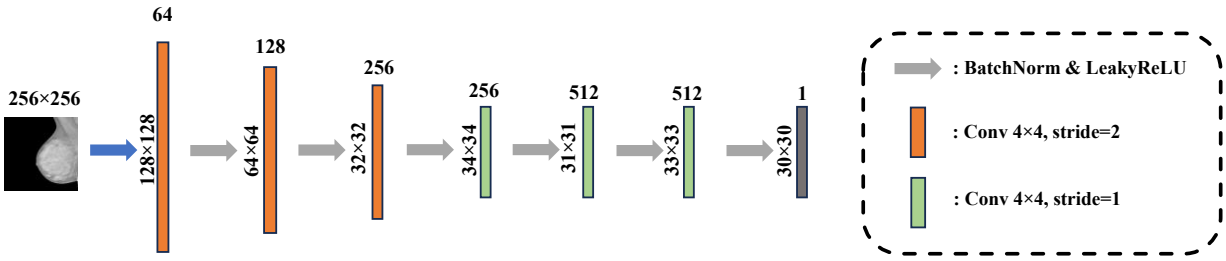

Figure 4. 256×256 resolution step's discriminator architecture.

The discriminator restricts the model to only focus on high-frequency structure details generation, while adopt L_1-loss to force low-frequency synthesis. The conditional adversarial loss and L_1-loss of EPG-GAN are formulated as follow:

$$L_{EPG-GAN}(G,D)=E_{x,y}[\log D(x,y)]+E_{x,G(x)}[\log(1-D(x,G(x)))] \tag{2}$$

$$L_{L1}(G)=\|y-G(x)\|_1 \tag{3}$$

$$G^*=\arg\min_G\max_D L_{EPG-GAN}(G,D)+L_{L1}(G) \tag{4}$$

G, D represent the generator and discriminator, respectively. x denotes the edge-guided mammogram input. y denotes the real mammograms.

3. Experiments and results

3.1. Dataset

In this study, we evaluated the proposed EPG-GAN on INbreast [19], which is a public database of FFDM images and prepared in DICOM format that take by many researchers. It presents 410 mammograms where several types of lesions (masses, calcifications, asymmetries, and distortions) were included of 115 unique patients. INbreast has the carefully associated ground truth annotation. Most of the databases, such as MIAS [20], only provide a circle around the area of interest [19]. The raw images were annotated with experts, and have an average size of 3328×3328 pixels.

To make the dataset better fit to our model, we first convert the DICOM format into PNG and resized the images to 512×512 pixels. Then, the images were resampled to the resolutions of 4×4, 128×128, and 256×256, for comparison of the synthetic performance under different resolution. Finally, the pixel intensities were also normalized.

3.2. Implementation details

In each experiment, a 24 GB Nvidia 4090 GPU was offered with Python version 3.11.5 and Pytorch version 2.1.1. The experimentation carried out in this research demonstrates how the proposed EPG-GAN is trained from scratch until good performance of image generation is achieved. The Adam optimizer was used with $\beta_1=0.0$, $\beta_2=0.99$, and the learning rate is 0.001. In order to fully use the GPU, we adapt different batch sizes for different resolution images. For 128×128 and 256×256 images, the batch size

is 8. For 512×512 images, the batch size is 2. The counterparts of the EPG-GAN model were trained with the hyperparameter setting in their origin paper.

3.3. Evaluation metrics

The experimental results of EPG-GAN and its counterparts were measured by Fréchet inception distance (FID), Peak signal-to-noise ratio (PSNR), and Structural similarity index (SSIM).

The FID is a metric used to measure the difference between an image generated by a generative model and a real image, and is computed by comparing the distances of their distributions in the feature space of the Inception v3 [49] model. The feature vector used in FID is a high-dimensional vector outputted from the penultimate fully-connected layer of the Inception v3 model, which captures the visual characteristics of the image.

$$FID = \left\| \mu_r - \mu_g \right\|^2 + Tr(\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{\frac{1}{2}}) \tag{5}$$

μ_r and μ_g are the mean vector of the real image and the generated image. Σ_r and Σ_g are the covariance matrix of the real image and the generate image. It is expected that a synthesized image with good quality should result in a lower FID score; otherwise, the image is assumed to be less similar to the real image. The highest similarity the two images may demonstrate is having an FID score of 0.

The PSNR measures the quality of synthesized image compared with the real image by using MSE. PSNR can provide information about the overall quality of the image.

$$PSNR(r,g) = 10\log_{10}(\frac{255^2}{MSE(r,g)}) \tag{6}$$

r and g are real image and generated image. A higher value of PSNR represented more acceptable quality of the image generated by the model.

The SSIM evaluates an image by considering it as multiple regions and comparing the structural, luminance, and contrast metrics within these regions.

$$SSIM(r,g) = \frac{(2\mu_r\mu_g + c_1)(2\sigma_{rg} + c_2)}{(\mu_r{}^2 + \mu_g{}^2 + c_1)(\sigma_r{}^2 + \sigma_g{}^2 + c_2)} \tag{7}$$

μ_r and μ_g are the mean vector of the real image and the generated image. σ_r^2 and σ_g^2 are the variance of the real image and the generated image. σ_{rg} is the covariance of the real image and the generate image. c_1 and c_2 are the constants. A higher value of SSIM represented more acceptable quality of the image generated by the model.

3.4. Results of EPG-GAN

In Table 1, several evaluation criteria were used to assess the experimental result in same resolution and different resolution to indicate the overall performance of EPG-GAN and its counterparts.

In the synthesis of 512×512 resolution mammograms process, the EPG-GAN achieves an average of 22.23, 34.89, and 0.983 for FID, PSNR, and SSIM. Meanwhile, the average performance of these indicators during the synthesis of 256×256 resolution mammograms process are 27.42, 34.26, and 0.974. In first step that the synthesis of 1 128×128 resolution mammograms process, EPG-GAN achieved 22.45, 35.18, and 0.976 for FID, PSNR, and SSIM. The results of EPG-GAN are significantly better than the early work by Oyelade *et al.* [41] (with 0.800 and 27.72 at SSIM and PSNR in the

synthesis of 256×256 resolution mammograms process), and EPG-GAN is validated on synthesis high-quality of 512×512 resolution mammograms compared with their 256×256 resolution.

In clinical diagnosis, extremely dense mammogram and heterogeneously dense mammogram are two types of mammograms that doctors and deep-learning model can easily make mistake due to the more gland area and less fat area overlapped the lesion area in the breast. However, these two types of mammograms generally get poor-labeled in mammogram datasets. It's critical to synthesis high-quality extremely dense mammogram and heterogeneously dense mammogram to augment the datasets. As shown in Figure 5, the results indicate that the EPG-GAN has achieved encouraging results on the mammogram images. Four typical cases are presented with a low-resolution synthesis mammogram (128×128), a high-resolution synthesis mammogram (512×512), and a real mammogram (512×512). The light (white) area in each mammogram is the gland area, it takes most of the whole breast. Our EPG-GAN firstly learned global feature in low-resolution mammogram synthesis process to ensure the accuracy of the main structure and focus on textural details in high-resolution.

The distribution of the values for the ten (10) samples of synthesized images are plotted in Figure 6, respectively, for PSNR and SSIM. As shown in Figure 6, Due to we used the edge information of the background texture as auxiliary sketch guidance to achieve high-fidelity synthesis, the value of SSIM has been stable for the ten (10) samples of synthesized images. That's indicate that the EPG-GAN has achieved encouraging results on the mammogram images.

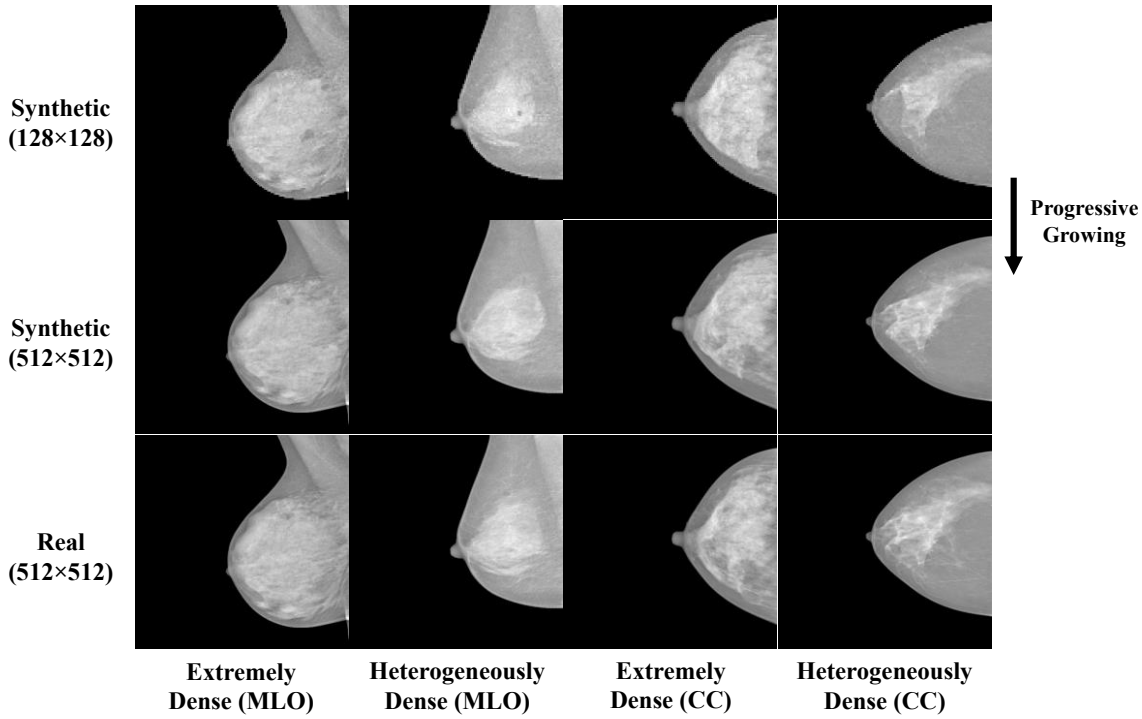

Figure 5. Qualitative comparison of synthesis mammogram with real mammogram.

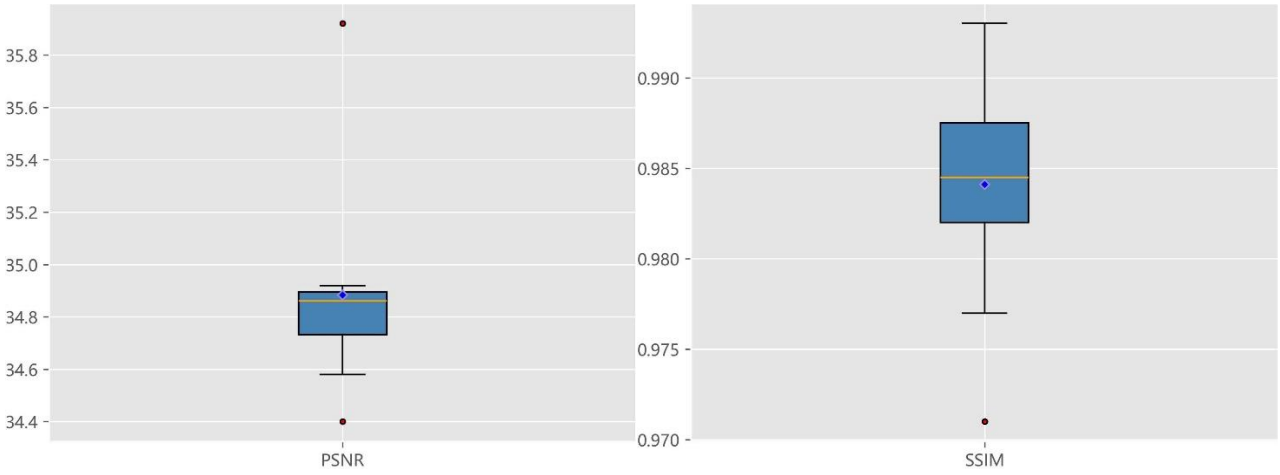

Figure 6. Boxplot showing the distribution of values obtained for ten randomly selected samples of synthesized images in computational metrics PSNR, SSIM.

3.5. Results of comparison models

Due to the limited work of this field, we evaluated the proposed EPG-GAN with three widely applied synthesis models including: MedGAN [50], cGAN [33] and RadiomicsFill-MET [51]. Moreover, MedGAN is a style-transfer model in 2020, cGAN is a conditional GAN-based synthesis models in 2014 and RadiomicsFill-MET is a diffusion-based synthesis model in 2024.

In Table 1, style-transfer-based model achieve a higher SSIM than other comparison models but slightly lower than EPG-GAN. The MedGAN model achieve approximate FID score at 23.81 (+1.58), slightly lower SSIM score at 0.916 (−0.067) and significantly lower PSNR score at 24.62 (−10.27) compared with EPG-GAN. The conditional GAN-based synthesis model counterpart cGAN achieves 25.29 higher FID score and 11.24 lower PSNR score than EPG-GAN, it shows obvious a lower ability in synthesis high-resolution mammograms. Furthermore, diffusion-based synthesis model RadiomicsFill-MET achieves significantly better in FID score (−22.17) and slightly lower PSNR score (−2.09) with much lower results in SSIM score (−0.129), it shows an approximate ability in synthesis high-resolution mammograms with EPG-GAN. The EPG-GAN outperforms the rest three counterparts by higher results in PSNR score and SSIM score. The results indicate that the proposed EPG-GAN has better robustness under high-resolution conditions, while many counterparts are getting lower performance as the resolution growing up.

Table 1. Quantitative comparison of performance of EPG-GAN and other methods on INbreast.

Method	Resolution	FID ↓	PSNR ↑	SSIM ↑
MedGAN [50]	512×512	23.81	24.62	0.916
cGAN [33]	512×512	47.52	23.65	0.896
RadiomicsFill-MET [51]	512×512	0.06	32.80	0.854
EPG-GAN	128×128	22.45	35.18	0.976
EPG-GAN	256×256	27.42	34.26	0.974
EPG-GAN	512×512	22.23	34.89	0.983

3.6. The effectiveness of edge-guided

To evaluate the edge-guided improvement to proposed EPG-GAN, an ablation study was drawn out under the same training and validating setting. Firstly, to evaluate the effectiveness of the edge-guided, we designed the Guided1, Guided2 models. In the Guided1 model, edge-guided image is not used as input. It is very difficult to generate high-resolution images directly in this case, so we reduced the target resolution of the generated image to 4×4 and adapt a simple backbone in this model. Meanwhile, we used a random vector as the new input. Therefore, we designed the Guided2 to compare the edge-guided in terms of the performance of the generated image.

As shown in Figure 7, Guided1 model takes a randomly initialized vector as input in epoch 1. The data distribution of this new input is completely different from that of the real image, so it is very time consuming for the generator searching through the whole hidden space of the real images to find a distribution that satisfies the discriminator. From epoch 1 to epoch 200, the generator explores the hidden space and eventually finds a data distribution that is approximately to the color distribution of the real image. But the generated image is totally not the same as the real image. The generator adjusting the image details and structure in order to be similar to the real image during the epoch 300 to epoch 500. After a long time of training, the generated image is closer to the real image. The whole process is much more time-consuming and ineffective than the process has edge-guided.

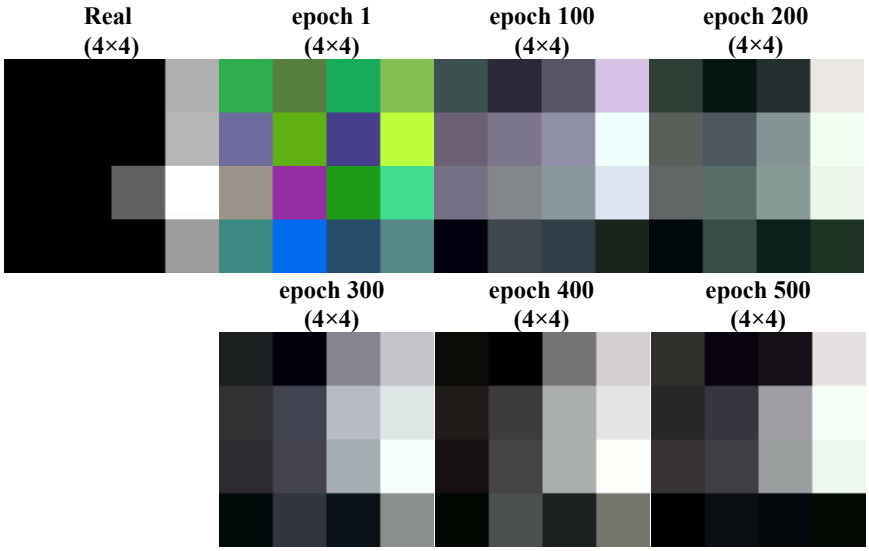

Figure 7. Synthetic images (4×4 resolution) without edge-guided.

In Table 2, the results indicate the edge-guided design of the EPG-GAN (Guided 2) model achieves evidently higher performance in all three criteria. The Guided2 model reaches the lowest FID score and the highest SSIM score with 18.99 and 0.896, which is significantly better than the Guided1. These results indicate that the edge-guided design is effective.

Table 2. Performance of edge-guided in synthetic images (4×4 resolution).

Model	FID ↓	PSNR ↑	SSIM ↑
Guided1	59.67	11.19	0.369
Guided2	18.99	12.28	0.896

3.7. Editable synthesis

Furthermore, we test our model for editable image synthesis. Via customized editing of label maps, our EPG-GAN can be used as a convenient tool to simulate various, meaningful mammograms for data augmentation. Compared with traditional data augmentation such as image rotation, our GAN-based augmentation method can provide greater variability with editable operations and therefore has great potential to improve performance.

Figure 8 shows synthesized example after editing the edge-guided label maps, using the proposed EPG-GAN. In INbreast, the early detection of breast cancer (benign or malignant) depends on the morphology and the location of lesion region. By adding or changing the position of the lesion region in the label map of a normal case, we can create a new case of mass with high fidelity.

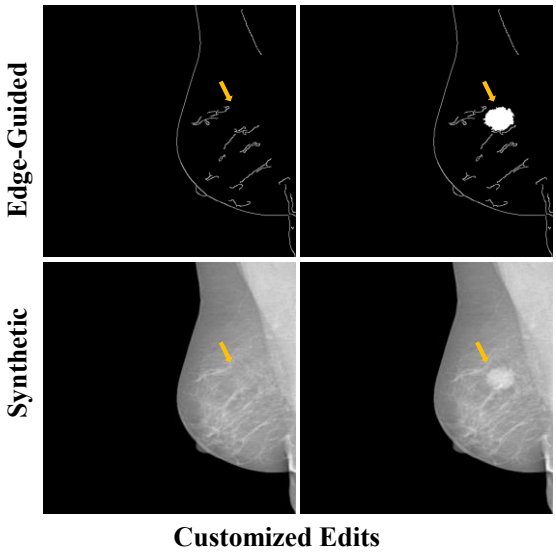

Figure 8. Example of editable synthesis. The yellow arrows indicate the changes before and after editing.

4. Conclusion

Through the edge-guided design, the proposed EPG-GAN can generate mammograms effectively and avoid synthesizing unnatural images in which the edge has blurred areas and distorted structures. Together with the progressive growing strategy, the introduction of the fusion block improves the ability that fade the image trained by previous step in the next step smoothly to gradually generate high-resolution mammograms.

Compared with the typical models, EPG-GAN achieves significantly higher PSNR score and SSIM score. At the starting stages, the EPG-GAN model leverages the strength of the edge-guided, which contributes to efficiently generate mammograms under complex surroundings. Benefiting from the progressive growing design, by introducing the fusion block, EPG-GAN fade the image trained by previous step in the next step smoothly and make sure that EPG-GAN can gradually generate mammograms from low to high resolution. Compared with the pure GAN-based models, EPG-GAN performs higher overall results in synthesis high-resolution mammograms.

The limitation of EPG-GAN is that our model only been tested on INbreast. Testing on other datasets with larger scale or different breast densities could be more accurate generalization performance. Additionally, the performance on downstream tasks (lesion classification or detection) in breast cancer diagnosis is equally important after augment the mammogram datasets. We need to fully test the

performance of the enhanced mammogram datasets in these downstream tasks. Finally, the balance between contrast, signal-to-noise ratio (SNR), and radiation dose for the synthesis mammograms is important, we need to fully test the property of the images. These are our future works to improve our model.

Furthermore, GAN-based method also shows great application potential in synthesis STEM images recently. Khan *et al.* [52] construct a Cycle-GAN model to synthesis realistic experimental STEM images by converting simulated STEM images that achieved the FID score at 0.35. Compered with Khan's model, EPG-GAN can only use experimental STEM images to achieve the same result theoretically. Due to the lack of STEM images dataset of our group, it is difficult for us to test the performance of EPG-GAN in material. We will test the performance of our method application potential in different fields after collecting the dataset.

In conclusion, the proposed EPG-GAN model achieves the best performance in synthesis high-resolution mammograms via its unique edge-guided progressive growing structure which can effectively combine the advantages of edge-guided and progressive growing. The EPG-GAN model can help to explore the high-resolution mammogram dataset augmentation technique and expand its potential for wider application.

Acknowledgments

The authors received no financial support for this article's research, authorship, and publication.

Conflicts of interests

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

Authors' contribution

Conceptualization, S.W. and S.W.; methodology, S.W.; software, S.W.; validation, S.W.; formal analysis, S.W.; investigation, S.W.; resources, S.W. and S.W.; data curation, S.W.; writing—original draft preparation, S.W. and S.W.; writing—review and editing, S.W. and S.W.; visualization, S.W.; supervision, S.W.; project administration, S.W.; funding acquisition, S.W. All authors have read and agreed to the published version of the manuscript.

References

[1] Bahri Y, Dyer E, Kaplan J, Lee J, Sharma U. Explaining scaling laws of neural network generalization. *arXiv* 2021, arXiv:2102.06701.

[2] Hestness J, Narang S, Ardalani N, Diamos G, Jun H, *et al.* Deep learning scaling is predictable, empirically. *arXiv* 2017, arXiv:112.00409.

[3] Kaplan J, McCandlish S, Henighan T, Brown TB, Chess B, *et al.* Scaling laws for neural language models. *arXiv* 2020, arXiv:2001.08361.

[4] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*. 2020, 70(1):7–30.

[5] Ginsburg O, Yip CH, Brooks A, Cabanes, A, Caleffi M, *et al.* Breast cancer early detection: a phased approach to implementation. *Cancer* 2020, 126(10):2379–2393.

[6] Peppard HR, Nicholson BE, Rochman CM, Merchant JK, Mayo RC, *et al.* Digital breast tomosynthesis in the diagnostic setting: indications and clinical applications. *Radiographics* 2015,

35(4):975–990.

[7] Nguyen T, Levy G, Poncelet E, Le Thanh T, Prolongeau J, *et al.* Overview of digital breast tomosynthesis: clinical cases, benefits and disadvantages. *Diagn. Interv. Imaging* 2015, 96(9):843–859.

[8] Eklund GW. The art of mammographic positioning. In *Radiological Diagnosis of Breast Diseases*. Berlin: Springer, 2000, pp.75–88.

[9] Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Comput. Med. Imaging and Graphic* 2007, 31(4–5):198–211.

[10] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. *J. Big Data* 2019, 6(1):1–48.

[11] Cheng J, Han Z, Mehra R, Shao W, Cheng M, *et al.* Computational analysis of pathological images enables a better diagnosis of TFE3 Xp11.2 translocation renal cell carcinoma. *Nat. Com.* 2020, 11:1–9.

[12] Lotter W, Diab AR, Haslam B, Kim JG, Grisot G, *et al.* Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nat. Med.* 2021, 27(2):244–249.

[13] Carneiro G, Nascimento J, Bradley AP. Automated analysis of unregistered multi-view mammograms with deep learning. *IEEE Trans. Med. Imaging* 2017, 36 (11):2355–2365.

[14] Dhungel N, Carneiro G, Bradley AP. A deep learning approach for the analysis of masses in mammograms with minimal user intervention. *Med. Image Anal*. 2017, 37:114–128.

[15] Lei Y, Tian Y, Shan H, Zhang J, Wang G. Shape and margin aware lung nodule classification in low-dose CT images via soft activation mapping. *Med. Image Anal*. 2020, 60:101628.

[16] Shen Y, Wu N, Phang J, Park J, Liu K, *et al.* An interpretable classifier for high-resolution breast cancer screening images utilizing weakly supervised localization. *Med. Image Anal*. 2021, 68:101908.

[17] Yang Z, Cao Z, Zhang Y, Tang Y, Lin X, *et al.* MommiNet-v2: Mammographic multi-view mass identification networks. *Med. Image Anal*. 2021, 73:102204.

[18] Sawyer-Lee R, Gimenez F, Hoogi A, Rubin D. Curated breast imaging subset of digital database for screening mammography (CBIS-DDSM). The Cancer Imaging Archive. 2016. Available: https://doi.org/10.7937/K9/TCIA.2016.7O02S9CY (accessed on 23 February 2025).

[19] Moreira IC, Amaral I, Domingues I, Cardoso A, Cardoso MJ. INbreast: toward a full-field digital mammographic database. *Acad Radiol*. 2012, 19(2):236–248.

[20] Suckling J. The mammographic image analysis society digital mammogram database. *Digit. Mammo*. 1994, 375–386.

[21] Cheng JZ, Ni D, Chou YH, Qin J, Tiu CM, *et al.* Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. *Scientific reports.* 2016, 6:1–13.

[22] Mark D, Lucy MW, Dominic W, Emma L, Alistair M, *et al.* OPTIMAM mammography image database: a large-scale resource of mammography images and clinical data. *arXiv* 2020, arXiv:2004.04742.

[23] Moein, Yue L, Hossein A, Edward A, Karin D, *et al.* CSAW-M: an ordinal classification dataset for benchmarking mammographic masking of cancer. *arXiv* 2021, arXiv:2112.01330.

[24] Jeong JJ, Vey BL, Bhimireddy A, Kim T, Santos T, *et al*. The EMory BrEast imaging Dataset

(EMBED): a racially diverse, granular dataset of 3.4 million screening and diagnostic mammographic images. *Radiology: Artificial Intelligence* 2023, 5(1):e220047.

[25] Liu F. Susan: segment unannotated image structure using adversarial network. *Magnetic Resonance in Med.* 2019, 81(5):3330–3345.

[26] Gao J, Jiang Q, Zhou B, Chen D. Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: an overview. *Math. Bio. and Eng*. 2019,16:6536.

[27] Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, *et al.* Gan-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing* 2018, 321:321–331.

[28] Salehi P, Chalechale A, Taghizadeh M. Generative adversarial net works (gans): an overview of theoretical model, evaluation metrics, and recent developments. *arXiv* 2020, arXiv:2005.13178.

[29] Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med. Image Anal.* 2019, 58:101552.

[30] Badano A, Graff CG, Badal A. Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in silico imaging trial. *JAMA Netw. Open* 2018, 1(7):e185474.

[31] Wen-mei W Hwu. GPU computing gems emerald edition. San Francisco: Morgan Kaufmann, 2011, pp.813–829.

[32] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, *et al.* Generative adversarial networks. *arXiv* 2014, arXiv:1406.2661.

[33] Mirza M, Osindero S. Conditional generative adversarial nets. *arXiv* 2014, arXiv:1411.1784.

[34] Salehinejad H, Valaee S, Dowdell T, Colak E, Barfett J. Generalization of deep neural networks for chest pathology classification in x-rays using generative adversarial networks. In *Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing*, Calgary, Canada, April 15–20, 2018, pp. 990–994.

[35] Frid-Adar M, Klang E, Amitai M, Goldberger J, Greenspan H. Synthetic data augmentation using gan for improved liver lesion classification. In *Proceedings of the IEEE 15th international symposium on biomedical imaging*, Washington, USA, April 4–7, 2018, pp. 289–293.

[36] Nie D, Trullo R, Lian J, Petitjean C, Ruan S, *et al.* Medical image synthesis with context-aware generative adversarial networks. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Quebec City, Canada, September 11–13, 2017, pp. 417–425.

[37] Wu E, Wu K, Cox D, Lotter W. Conditional infilling gans for data augmentation in mammogram classification. In *Proceedings of the Image Analysis for Moving Organ, Breast, and Thoracic Images*, Granada, Spain, September 16 and 20, 2018, pp. 98–106.

[38] Jin D, Xu Z, Tang Y, Harrison AP, Mollura DJ. Ct-realistic lung nodule simulation from 3d conditional generative adversarial networks for robust lung segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Granada, Spain, September 16–20, 2018, pp. 732–740.

[39] Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review. *Med. Image Anal.* 2019, 58:101552.

[40] Shen T, Hao K, Gou C, Wang FY. Mass image synthesis in mammogram with contextual information based on GANs. *Comput. Methods Programs Biomed*. 2021, 202:106019.

[41] Oyelade ON, Ezugwu AE, Almutairi MS, Saha AK, Abualigah L. A generative adversarial network for synthetization of regions of interest based on digital mammograms. *Sci. Rep*. 2022, 12(1):6166.
[42] Joseph AJ, Dwivedi P, Joseph J, Francis S, Pournami PN, *et al.* Prior-guided generative adversarial network for mammogram synthesis. *Biomed. Signal Process. Control* 2024, 87:105456.
[43] Shin Y, Qadir H A. Abnormal colon polyp image synthesis using conditional adversarial networks for improved detection performance. *IEEE Access*. 2018, 6:56007–56017.
[44] Zhang T, Fu H. Skrgan: sketching-rendering unconditional generative adversarial networks for medical image synthesis. *arXiv* 2019, arXiv:1908.04346.
[45] Karras T, Aila T, Laine S, Lehtinen J. Progressive growing of gans for improved quality, stability, and variation. *arXiv* 2017, arXiv:1710.10196.
[46] Canny J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 1986, PAMI-8(6):679–698.
[47] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *arXiv* 2015, arXiv:1505.04597.
[48] Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. *arXiv* 2016, arXiv:1611.07004.
[49] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. *arXiv* 2015, arXiv:1512.00567.
[50] Armanious K, Jiang C, Fischer M, Küstner T, Hepp T, *et al.* MedGAN: medical image translation using GANs. *arXiv* 2018, arXiv:1806.06397.
[51] Na I, Kim J, Ko ES, Park H. Radiomics Fill-Mammo: synthetic mammogram mass manipulation with radiomics features. *arXiv* 2024, arXiv:2407.05683.
[52] Khan A, Lee CH, Huang PY, Clark BK. Leveraging generative adversarial networks to create realistic scanning transmission electron microscopy images. *npj Comput. Mater.* 2023, 9(1):85.