

Research Article

Reliable Breast Cancer Diagnosis with Deep Learning: DCGAN-Driven Mammogram Synthesis and Validity Assessment

Dilawar Shah ^[],¹ Mohammad Asmat Ullah Khan,¹ and Mohammad Abrar²

¹Department of Computer Science, Faculty of Computing and Information Technology, International Islamic University, Islamabad 44000, Pakistan

²Department of Computer Science, Bacha Khan University, Charsadda 24420, Pakistan

Correspondence should be addressed to Dilawar Shah; dilawar.phdcs151@iiu.edu.pk

Received 29 September 2023; Revised 8 January 2024; Accepted 2 February 2024; Published 19 February 2024

Academic Editor: Soumya Ranjan Nayak

Copyright © 2024 Dilawar Shah et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Breast cancer imaging is paramount to quickly detecting and accurately evaluating the disease. The scarcity of annotated mammogram data presents a significant obstacle when building deep learning models that can produce reliable outcomes. This paper proposes a novel approach that utilizes deep convolutional generative adversarial networks (DCGANs) to effectively tackle the issue of limited data availability. The main goal is to produce synthetic mammograms that accurately reproduce the intrinsic patterns observed in real data, enhancing the current dataset. The proposed synthesis method is supported by thorough experimentation, demonstrating its ability to reproduce diverse viewpoints of the breast accurately. A mean similarity assessment with a standard deviation was performed to evaluate the credibility of the synthesized images and establish the clinical significance of the data obtained. A thorough evaluation of the uniformity within each class was conducted, and any deviations from each class's mean values were measured. Including outlier removal using a specified threshold is a crucial process element. This procedure improves the accuracy level of each image cluster and strengthens the synthetic dataset's general dependability. The visualization of the class clustering results highlights the alignment between the produced images and the inherent distribution of the data. After removing outliers, distinct and consistent clusters of homogeneous data points were observed. The proposed similarity assessment demonstrates noteworthy effectiveness, eliminating redundant and dissimilar images from all classes. Specifically, there are 505 instances in the normal class, 495 instances in the benign class, and 490 instances in the malignant class out of 600 synthetic mammograms for each class. To check the further validity of the proposed model, human experts visually inspected and validated synthetic images. This highlights the effectiveness of our methodology in identifying substantial outliers.

1. Introduction

Breast cancer (BC) is a popular type of cancer in females that forms in the cells of the breast. It is triggered by the abnormal cell's growth in the breast and divides uncontrollably, forming a tumor. Its early detection is crucial, as it allows for timely treatment, thus saving many lives [1]. In the year 2020, a total of 2.26 million new instances of breast cancer were identified globally, accounting for approximately 11.7% of the overall incidence of cancer. Currently, this form of cancer has become the most prevalent, surpassing lung cancer with a prevalence rate of 11.4%. Certain regions, namely, Australia, Europa (North, West, and South), and North America, appear to be more significantly impacted than others [2]. Moreover, according to the World Health Organization [3], the incidence of breast cancer is projected to rise from 2.26 million new cases in 2020 to 3.19 million in 2040, reflecting a notable rising trend of 41% over the following two decades. Cancer is becoming more prevalent in Pakistan, with 19 million new cancers of all types recorded in 2020. In the context of Pakistan, the annual incidence of breast cancer is predicted to exceed 83,000 cases. Annually, over 40,000 women succumb to this debilitating ailment [4]. Medical practitioners use several modalities for early prediction of breast cancer, like mammography, ultrasound, and magnetic resonance imaging (MRI). Mammography is a low-dose X-ray of the breast and is the most common screening tool for breast cancer [5]. However, accurately interpreting mammograms is challenging as these images are complex, and detecting abnormalities can be difficult. Human reading of mammograms can sometimes produce false-positive or negative results. Variability in expertise can also lead to inconsistencies in diagnosis and treatment decisions [6].

The artificial intelligence community is trying to help with breast cancer detection. Training algorithms on a large mammogram and clinical datasets conducts extensive research. These models can recognize breast cancer patterns and accurately analyze mammograms, reducing false positives and negatives and helping radiologists make better diagnoses and treatment decisions. Machine learning (ML) models can standardize analysis and reduce radiologists' expertise [7, 8]. Machine learning algorithms can identify subtle patterns and features that may not be readily noticeable to human observers and can aid in the early detection of BC.

ML can identify breast cancer, but it has limits. Training data are crucial to such models. The model's performance may be limited if the training data doesn't include unusual or atypical breast cancer cases. Comprehensive and representative datasets are essential to detecting breast cancer across populations and variations. These models may fail to generalize to new data that differs significantly from the training data [9–11].

Researchers [12] increasingly acknowledge the significance of tackling the issue of limited data representation to enhance the performance and generalizability of these models. The technique of inflating existing datasets is widely employed. Data augmentation is a methodology that encompasses the generation of novel training samples by applying diverse transformations to the preexisting data. The augmentation of mammogram images can encompass various techniques, such as flipping, rotation, scaling, or introducing noise. These techniques serve the purpose of enhancing the diversity and variability of the training data. This feature enables the model to acquire knowledge from a more extensive array of instances. However, this approach also has limitations.

In conjunction with continuous research and development endeavors, these methodologies strive to tackle the issue of inadequate data representation in machine learning and enhance the efficacy and applicability of models across diverse fields [13–16].

The primary objective of this study is to employ a DCGAN for data augmentation and subsequently validate its effectiveness. DCGANs can create synthetic mammograms from minimal real data, as shown in Figure 1.

The primary motivation is to overcome data shortages, enhance diagnostic precision, and advance breast cancer imaging using generative models to synthesize realistic mammograms. The generated data could improve deep learning models and aid in providing patients with better care. The study aspires to improve patients' lives worldwide by improving such models' diagnostic accuracy through more robust data augmentation. The accuracy of the proposed models helps clinicians detect breast cancer earlier. The study's objectives include utilizing DCGANs to create synthetic mammograms to fill the data gap and increase the dataset's diversity and clinical applicability. Enhancing deep learning by exposing it to a more extensive and varied dataset during training can improve the precision of deep learning models for detecting breast cancer. Increasing the number of cases available for comparison and validation through synthetic images will boost radiologists' confidence in their diagnostic abilities.

Section 2 presents recent literature on GAN. Section 3 presents the proposed methodology for mammogram augmentation and subsequent validation. The results of the study are presented in Section 4. The conclusion and future work are discussed in Section 5.

2. Related Work

In the context of mammography, several data augmentation approaches are frequently employed to improve the performance and generalizability of the models. Machine learning has experienced substantial progress in the last decade, primarily due to advancements in deep neural networks. These networks have demonstrated exceptional performance in several medical imaging tasks, contributing to the increased popularity of machine learning in this domain. Meanwhile, the generative modelling and data synthesis field has made significant advancements in quality, mainly attributed to the emergence of generative adversarial networks. GANs currently exhibit remarkable capabilities for generating visually realistic images that closely resemble the content of the datasets they were trained on.

Wu et al. [18] conducted research to address the data scarcity and imbalance problem in breast cancer detection using a publicly available dataset from the UK, namely the OPTIMAM Mammography Image Database. The dataset contains 8282 malignant, 1287 benign, and 16887 normal mammographic images. They divided the data into 60% for training, 20% for validation, and 20% for testing. They trained a contextual GAN model to augment the dataset with the self-attention mechanism. They used both traditional and GAN-based augmentations. Their GAN-augmented model produced an AUC of 0.846. The model performs a binary classification of normal and malignant.

Desai et al. [19] developed the DCGAN model using a benchmark DDSM dataset. They utilized 218 for training and 47 each for testing and validation. Their experiments reported an accuracy of 78.23% when the model was trained on original images. At the same time, the combination of synthetic and authentic images produced an enhanced accuracy of 87% with an improvement factor of 8.77. The authors show that GAN is a workable choice for training such models with a data shortage.

Alyafi et al. developed a DCGAN model for breast mass augmentation using a subset of 80000 images from the UKbased OPTIMAM mammography image database (OMI-DB) [17]. The authors demonstrate the performance of a classifier in an imbalanced dataset with and without synthetic data in the experiments. They created breast mass patches with 128×128 pixel dimensions using a modified



(a)

(b)

FIGURE 1: (a) GAN-generated images (b) real images [17].

version of DCGAN. GAN augmentation was compared to traditional augmentation. The results show that using DCGANs with flipping augmentation improves the *F*1 score by up to 0.09 compared to the original mammographic images. The job can be expanded to include other similar tasks. Their work is limited to small mammogram patches.

Shen et al. [20] developed a GAN-based system using a benchmark DDSM dataset and a local dataset collected from Nanfang Hospital, China. The study aimed to address the issue of limited data in medical image analysis by designing a model to generate labelled images based on contextual information within the breast mammograms. The model was evaluated, and the results showed that their augmentation technique increased the diversity of the dataset and achieved an improvement of 5.03% in the detection rate. The model is a viable option for generating labelled breast images.

In [21], they proposed a deep learning-based mammogram recognition model. The model performs a special autoencoder-generative adversarial network (AGAN) for data augmentation. The generator produces additional images in a perfect way for training the model. The final set of original and generated images is given as input to the CNN for classification. A total of 11,218 ROIs of mammograms from DDSM were used in the experiments. They reported an average accuracy in detecting abnormal vs healthy cases of 89.71%. The specificity was 80.58%, while the sensitivity and AUC were 93.54% and 0.9410, respectively. The work's main contribution was its novelty in its data augmentation compared to the other deep learning methods. The proposed model AGAN is learned only on normal data. The model does not consider other mammographic datasets.

Another study generated breast mammograms with GANs [22]. Their main aim was to detect mammographically occult (MO) cancer in women with dense breasts. The researchers employed a convolutional neural network (CNN). The network was trained on processed mammographic images from the Radon cumulative distribution transform (RCDT) 1366 processed mammograms collected from the University of Pittsburgh Medical Center, USA. They reported an AUC of 0.77. The system can identify patients for further screening in the early detection of MO-related cancer. However, they did not consider benchmark datasets.

The authors developed a StyleGAN 2 system using 105,948 normal mammograms collected from Asan Medical Center, Korea, from January 2008 to December 2017 [23]. They evaluated GAN-generated images through Fréchet Inception Distance (FID) equal to 4.383 and the Inception Score of 16.67. The multiscale structural similarity index measure (MS-SSIM) stood at 0.39, and the average value of the peak signal-to-noise ratio (PSNR) was 31.35. Their model has performed with reasonable fidelity to real images. The system was only limited to normal mammographic local images. The summary of the literature is presented in Table 1.

This study presents an innovative approach to addressing the scarcity of annotated mammogram data by employing DCGANs. This methodology is adept at generating synthetic mammograms that mirror real-data characteristics with high fidelity. Key contributions of this research are outlined as follows:

- (i) The research extensively tests the effectiveness of the DCGAN-based synthesis in accurately replicating various mammographic features, including diverse tissue types, lesion characteristics, and breast views. The quality and authenticity of these synthetic images are meticulously evaluated using mean similarity measures and standard deviation analyses, ensuring a rigorous assessment of their realism.
- (ii) The study employs a systematic approach to enhance data precision by identifying and removing outliers. This is achieved through a threshold-based outlier removal mechanism, significantly bolstering the synthetic dataset's reliability. The refined dataset demonstrates clinical relevance, as evidenced by its consistency across different classes.

References and year	Dataset	No. of samples	Technique	Key points	Limitation
Wu et al. [18]	Optimam mammography image Database	26456	Traditional + contextual GAN	The model performance on malignancy classification was improved	Classification of normal and malignant only
Desai et al. [19]	DDSM	287	DCGAN	The authors show that GAN is a workable choice for training such models with a data shortage	The model is evaluated with only two batch sizes
Alyafi et al. [17]	OPTIMAM mammography image Database (OMI-DB)	80000	DCGAN	The work can be extended to other similar tasks	Their work is limited to small patches of mammograms
Shen et al. [20]	DDSM and local dataset collected from Nanfang Hospital, China	11218	GAN	The model is a viable option for generating labelled breast images	Due to the inherent complexity and variability of breast tissue structures, this type of GAN might not be able to produce various realistic images
Swiderski et al. [21]	DDSM	11218	AGAN	The novelty of the model in data augmentation as compared to other deep learning models	Modified GAN is based on autoencoder architecture, which may misclassify key features for BC diagnosis
Lee et al. [22]	The local dataset was collected from the University of Pittsburgh Medical center, USA	1366	CGAN	The system can identify patients for further screening in the early detection of MO-related cancer	They did not consider benchmark datasets
Park et al. [23]	Local dataset from Asan medical center, Korea	105,948	StyleGAN2	Their model has comparable fidelity to real mammograms	The system was only limited to normal mammographic images

ă	
chniq	
Ţ	
Augmentation	2
Mammogram	
GAN-based	
of various	
omparison	
I: Č	



FIGURE 2: Proposed methodology at Glance.

- (iii) The reliability of the proposed model is further corroborated through visual validation conducted by expert radiologists. Their professional assessment confirms the clinical accuracy and utility of synthetic mammograms.
- (iv) The study showcases the consistency of the synthetic dataset through detailed visualizations of class clustering. These visualizations highlight the congruence between the generated mammograms and the real data distribution. The substantial number of images from each class passing the similarity assessment underscores the success of the proposed validation mechanism.

3. Methodology

This section describes the comprehensive methodology used to determine the reliability of the dataset and ensure the validity of the generated mammogram classes. The methodology includes creating mammogram classes with a DCGAN, determining similarity, and removing outliers using a three-fold standard deviation threshold. The overall methodology is depicted in Figure 2. This study aims to evaluate the generated classes' quality methodically and improve the dataset's robustness.

3.1. Data Collection. DDSM (digital database for screening mammography) is a benchmark dataset [24]. The DDSM dataset consists of a more extensive collection of 2,620 digital mammograms in DICOM format from 262 patients. The dataset consists of 695 normal mammograms, 141 benign mammograms without callback, 870 benign, and 914 malignant mammograms. The description of DDSM contains the ground truth information associated with each mammogram image with suspect lesions. It includes both benign and malignant cases, offering a diverse range of breast abnormalities.

3.2. Data Preparation. Preparing appropriate input data for the model is essential to ensuring consistency and clinical relevance. Medical images, including mammograms, can be susceptible to noise and artifacts that might affect the quality of the training data. A denoising algorithm, such as median filtering or wavelet denoising, is applied to the mammogram images. This denoising process effectively reduces noise while preserving diagnostically relevant features, resulting in cleaner images for training. To feed data to the network, the mammograms were resized into the same size and format.

3.3. DCGAN Architecture. The architecture of the DCGAN [25] plays a pivotal role in generating realistic mammogram images. This subsection presents a comprehensive overview of its architecture tailored to the mammogram generation task, including detailed descriptions and tables depicting key components. Its general architecture is shown in Figure 3.

Dotted arrows show fake mammograms. First, a noise batch z is generated; forward z through Generator (G); forward the real and fake batches through Discriminator D; calculate LD; update D; calculate LG; and update G.

In the diagram, random latent vector samples are taken from z Pz; Pz = N(0, 1) for each training iteration (see step 1 in the above diagram). After being normalized to the range [1, 1], this pure-noise batch is sent through G to create a set of fake images (G(z), step 2). As shown in step 3 with the dashed arrows, these fake images are normalized to the interval [0, 1] before passing through D to obtain realism probabilities. In step 4, LD is calculated, and D parameters are updated in step 5. After that, the fake batch is forwarded through D, and LG is calculated in step 6. Backpropagation is done eventually to update the parameters of G in step 7.

3.3.1. Generator Network. A random noise vector is fed into the generator network, which gradually converts it into synthetic mammogram images. It starts with convolutional layers and then adds nonlinearity with batch normalization and ReLU activation functions. Skip connections preserve key features during the downsampling process; they were inspired by U-Net architectures. The generator's architecture is summarized in Table 2.



FIGURE 3: General architecture of DCGAN.

Layer no	Layer type	Output shape	Activation
1	Dense (input)	$512 \times 4 \times 4$	_
2	Reshape	$256 \times 4 \times 4$	_
3	Conv2DTranspose	$512 \times 8 \times 8$	ReLU
4	Batch normalization	$256 \times 8 \times 8$	_
5	Conv2DTranspose	$128 \times 16 \times 16$	ReLU
6	Batch normalization	$64 \times 16 \times 16$	_
7	Conv2DTranspose	$32 \times 32 \times 32$	ReLU
8	Batch normalization	$32 \times 32 \times 32$	_
9	Conv2DTranspose	$3 \times 64 \times 64$	Tanh

3.3.2. Discriminator Network. The primary function of the discriminator network is to discern and differentiate between authentic mammogram images and artificially generated ones. The architecture consists of convolutional layers, followed by batch normalization and LeakyReLU activation functions to introduce nonlinearity. The discriminator's architecture is summarized in Table 3.

The training employs adversarial loss functions, such as binary cross-entropy or Wasserstein loss, to simultaneously optimize the generator and discriminator networks. Adam optimizer is utilized for its robustness in handling nonstationary data and complex loss landscapes.

3.4. Training Process. A crucial stage in this research is the network's training process, during which the generator learns to create realistic mammogram images, and the discriminator develops its capacity to tell real images from fake ones. The key components of the training process are described in this subsection, including the tuning of hyperparameters, loss functions, and convergence monitoring, as shown in Table 4.

The generator and discriminator networks compete in a two-player minimax game as part of the adversarial training approach. While the discriminator strives to become more accurate in distinguishing real from fake images, the generator seeks to reduce the discriminator's ability to differentiate between real and synthetic mammogram images.

TABLE 3: The discriminator architecture.

Layer no	Layer type	Output shape	Activation
1	Conv2D (input)	$64 \times 32 \times 32$	LeakyReLU
2	Conv2D	$64 \times 16 \times 16$	LeakyReLU
3	Batch normalization	$64 \times 16 \times 16$	
4	Conv2D	$128 \times 8 \times 8$	LeakyReLU
5	Batch normalization	$128 \times 8 \times 8$	
6	Conv2D	$256 \times 4 \times 4$	LeakyReLU
7	Batch normalization	$256 \times 4 \times 4$	—
8	Conv2D	$512 \times 4 \times 4$	_
9	Batch normalization	$512 \times 4 \times 4$	LeakyReLU
10	Flatten	$1 \times 1 \times 1$	Sigmoid

TABLE 4: Training hyperparameters.

Hyperparameter	Value
Learning rate (generator)	0.0002
Learning rate (discriminator)	0.0002
Batch size	64
Number of epochs	100
Optimizer	Adam
Beta 1 (Adam)	0.5
Beta 2 (Adam)	0.999

3.4.1. Loss Function. During training, the DCGAN uses binary cross-entropy loss as the primary loss function for the generator and discriminator. For the discriminator's real/ fake classification, this loss measures the difference between predicted and ground truth labels. The Wasserstein loss

```
Begin
  CosineSimilarity (A, B) {
    RETURN DotProduct (A, B)/(Norm (A) * Norm (B));
  CalculateMeanVector (Vectors) {
    RETURN average of vectors along each dimension;
  CalculateStandardDeviation (Vectors) {
    RETURN standard deviation of vectors along each dimension;
  CalculateCosineSimilarities (ImageVectors) {
    CosineSimilarities ← Empty list;
    FOR each vector in ImageVectors {
       Cosine \leftarrow CosineSimilarity (Vector, CalculateMeanVector (All previous vectors in ImageVectors));
       Append Cosine to CosineSimilarities;
    RETURN CosineSimilarities;
  CalculateThresholds (CosineSimilarities) {
    Mean ← calculate mean of CosineSimilarities;
    StdDev ← CalculateStandardDeviation (CosineSimilarities);
    ThresholdHigh \leftarrow mean + 3 * StdDev;
    ThresholdLow \leftarrow mean – 3 * StdDev;
    RETURN ThresholdHigh, ThresholdLow;
  FilterImagesByCosineSimilarity (ImageVectors) {
    CosineSimilarities ← CalculateCosineSimilarities (ImageVectors);
    ThresholdHigh, ThresholdLow ← CalculateThresholds (CosineSimilarities);
    FilteredImages ← Empty list;
    FOR each Cosine in CosineSimilarities {
       IF Cosine is within ThresholdHigh and ThresholdLow {
         Append corresponding ImageVector to FilteredImages;
    RETURN FilteredImages;
End
```

ALGORITHM 1:Pseudo-code for similarity assessment and outliers removal.

enhances gradient flow and stabilizes training during adversarial training.

3.5. Performance Evaluation. Assessing its output performance and quality is critical to understanding the network's effectiveness and clinical applicability. The evaluation metrics used to evaluate the synthetic mammograms thoroughly are presented in this subsection.

The outliers were identified and eliminated as necessary. The outliers were eliminated above or below the stipulated threshold. The proposed methodology is shown in the

following pseudo-code (see Algorithm 1). Combining these evaluation metrics ensures a robust and multidimensional assessment of the DCGAN-generated mammograms. By scrutinizing the synthetic images' structural, statistical, and diagnostic characteristics, the study gains valuable insights into the DCGAN's performance and its potential contribution to advancing medical imaging research and clinical practice.

4. Experiments and Results

This section presents the study's results on the application of DCGAN for mammogram generation and the subsequent validation process. The study focuses on three distinct classes of mammograms. It analyzes the mean similarity of each type and the distances of individual data points from their respective means using a statistical approach involving the three times standard deviation criterion.

4.1. Synthetic Mammogram Generation. The proposed network was first trained on images from three different classes of mammograms. The network was taught to make fake mammograms with many of the same features and characteristics as real ones. During training, the model learned each class's unique patterns and structures. This gave it the ability to make high-quality fake mammograms. The proposed model generated 600 images for each class during the entire training.



FIGURE 4: (a) Synthesized images from epoch 0 during training (b) synthesized images from epoch 1 during training.



FIGURE 5: (a) Synthesized images from epoch 2 during training (b) synthesized images from epoch 3 during training.

Figures 4(a) and 4(b) represent the images generated during the initial phases. Initially, the training process takes place over random noise.

Figure 5(a) represents the synthesized images from epoch 2 during training, while Figure 5(b) shows synthesized images from epoch 3 during the training process of the DCGAN.

Figure 6(a) shows synthesized images from epoch 45 during the training process, while Figure 6(b) shows synthesized images from epoch 50 during the training of DCGAN.

Figure 7(a) represents synthesized images from epoch 99 during the training of the DCGAN. The images are closer to the real ones. Figure 7(b) Shows final synthesized images from epoch 100 during training. These are the finest images of the proposed model during the entire training.

Figure 8 shows the losses of both the discriminator and generator networks. Figure 9 shows the real and fake images of the proposed model during training.

4.2. Mean Similarity Assessment. After generating synthetic mammograms, all synthetic and original images were mixed class-wise. For each class, the mean similarity is calculated, which provides insight into the consistency and similarity of the generated mammograms within that class.

During validation of the synthetic images in a normal class, 95 images were declared outliers out of a total of 600 images. In the benign class, 105 images were used, while in the malignant class, 110 images were declared outliers as per the similarity score, as shown in Figures 10–12.

4.3. Statistical Validation. The distance of each data point from its respective class mean is computed further to validate the quality and authenticity of the generated mammograms. This distance calculation involved utilizing the three times standard deviation criterion, which enabled us to quantify how much each generated mammogram deviated



FIGURE 6: (a) Synthesized images from epoch 45 during training (b) synthesized images from epoch 50 during training.



FIGURE 7: (a) Synthesized images from epoch 99 during training (b) synthesized images from epoch 100 during training.



FIGURE 8: Discriminator and generator network losses during training.

from the class mean. A more considerable distance value indicated a higher level of dissimilarity, whereas a smaller value suggested a closer resemblance to the mean.

The statistical validation approach is assessed to bolster the credibility of the generated mammograms. This involved calculating the distance of each image from its class mean



FIGURE 9: Real vs. fake losses during training.



FIGURE 10: (a) Presents similarity score of normal class. (b) Presents similarity score with outliers in red of normal class. (c) Presents similarity score of normal class without outliers.



FIGURE 11: (a) Presents similarity score of benign class. (b) Presents similarity score with outliers in red of benign class. (c) Presents similarity score of benign class without outliers.

using the three times standard deviation criterion. More considerable distances indicated more significant dissimilarity, while smaller distances suggested better alignment with the class mean. In Figure 13, distinct and coherent clusters of similar data points were evident after removing outliers. This highlighted the effectiveness of the proposed approach in forming meaningful clusters. Distance-based validation



FIGURE 12: (a) Presents the similarity score of the malignant class. (b) Presents similarity score with outliers in red of malignant class. (c) Presents the similarity score of the malignant class without outliers.



FIGURE 13: (a) Show the similarity of data points of each cluster without outliers. (b) Show the similarity of data points of each cluster with outliers in red.

methods provided a robust means of quantifying the authenticity of the synthetic mammograms, improving accuracy, and strengthening the reliability of the generated data for breast cancer imaging applications.

4.4. Validation from Human Experts. Considering how realistic some of the DCGAN-generated images look, we asked three medical experts with more than 10 years of experience in radiology and mammography to classify synthetic and real images. Each radiologist was shown 80 images of a 50/50 mixture of real and synthetic images and was asked to rank them based only on their visual appreciation. The experts achieved an average accuracy of only 68%, thus showing how visually accurate the generated images are. In the Expert Panel Review phase, a group of radiologists with extensive experience in mammography evaluated the synthetic mammogram images. This panel was carefully selected based on their clinical expertise and familiarity with mammographic interpretation. They conducted a detailed assessment of each synthetic image, focusing on critical diagnostic features such as tissue density, lesion characterization, and calcifications or other anomalies indicative of potential pathology. Their assessment aimed to determine the realism and diagnostic accuracy of the synthetic images, comparing them to actual mammograms. The radiologists' feedback provided valuable insights into the clinical viability of the synthetic images, ensuring that they met the standards required for effective diagnostic use in a clinical setting.

5. Conclusions and Future Work

This research aims to give an in-depth study into the utilization of DCGAN for the generation of mammograms and the following validation. The research was centered on three classes of mammography, and it utilized a statistical methodology that involved the three times standard deviation criterion in examining the mean similarity of each class and the distances of individual data points from their respective means. The findings demonstrate that the proposed network can successfully generate synthetic mammograms that exhibit traits and properties comparable to real mammograms. As a result of rigorous training, the network was able to gain the capability to make synthetic images of high quality that capture the one-of-a-kind patterns and structures that are characteristic of each class. This was demonstrated by the synthetic images produced during the various training epochs. The calculation of mean similarity offered insights into the consistency and similarity of the generated mammograms within each class, further highlighting the network's capacity to capture class-specific properties. These insights are evident from the calculation of mean similarity. The statistical validation strategy relied on calculating distances between mammograms to ensure the generated mammograms were genuine. Some of the generated images were also validated by the human radiologist, confirming the authenticity of the proposed model. The research provided a reliable approach for evaluating dissimilarity and alignment by first estimating the level of variation from class means and then utilizing the three times standard deviation criterion as the measuring stick. Notably, eliminating outliers showed cohesive and distinct clusters of similar data points, confirming that the strategy effectively produces meaningful clusters. We plan to test more datasets with more GAN architectures in the future.

Data Availability

The dataset used in this study is publicly available at [24]. Code will be made available on request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- R. M. Mann, A. Athanasiou, P. A. T. Baltzer et al., "Breast cancer screening in women with extremely dense breasts recommendations of the European Society of Breast Imaging (EUSOBI)," *European Radiology*, vol. 32, no. 6, pp. 4036–4045, 2022.
- [2] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [3] F. De Bueger, M. Decroës, and B. Macq, "Deep learning in mammography: reducing annotation effort for breast mass segmentation," *Master in Mathematical Engineering, Ecole polytechnique de Louvain*, Université catholique de Louvain, Ottignies-Louvain-la-Neuve, Belgium, 2022.
- [4] S. International, "Importance of breast cancer awareness in Pakistan," *Shifa International Patients*, vol. 2023, 2022.
- [5] NIBIB, "Mammography," 2023, https://www.nibib.nih.gov/ science-education/science-topics/mammography.

- [6] N. Sharma, A. Y. Ng, J. J. James et al., "Multi-vendor evaluation of artificial intelligence as an independent reader for double reading in breast cancer screening on 275,900 mammograms," *BMC Cancer*, vol. 23, no. 1, pp. 460–513, 2023.
- [7] L. Abdelrahman, M. Al Ghamdi, F. Collado-Mesa, and M. Abdel-Mottaleb, "Convolutional neural networks for breast cancer detection in mammography: a survey," *Computers in Biology and Medicine*, vol. 131, Article ID 104248, 2021.
- [8] R. Kumar, "Double internal loop higher-order recurrent neural network-based adaptive control of the nonlinear dynamical system," *Soft Computing*, vol. 27, no. 22, pp. 17313– 17331, 2023.
- [9] I. Zyout, "Evaluation of deep convolutional neural networks for detecting nonpalpable breast abnormalities in mammography," in 2023 Advances in Science and Engineering Technology International Conferences (ASET), pp. 1–6, IEEE, Dubai, UAE, February 2023.
- [10] R. Kumar, "Memory recurrent Elman neural network-based identification of time-delayed nonlinear dynamical system," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 2, pp. 753–762, 2023.
- [11] R. Kumar, S. Srivastava, and J. Gupta, "Modeling and adaptive control of nonlinear dynamical systems using radial basis function network," *Soft Computing*, vol. 21, no. 15, pp. 4447–4463, 2017.
- [12] P. Oza, P. Sharma, S. Patel, F. Adedoyin, and A. Bruno, "Image augmentation techniques for mammogram analysis," *Journal of Imaging*, vol. 8, no. 5, p. 141, 2022.
- [13] R. Kumar, S. Srivastava, J. Gupta, and A. Mohindru, "Selfrecurrent wavelet neural network-based identification and adaptive predictive control of nonlinear dynamical systems," *International Journal of Adaptive Control and Signal Processing*, vol. 32, no. 9, pp. 1326–1358, 2018.
- [14] R. Kumar and S. Srivastava, "Externally Recurrent Neural Network based identification of dynamic systems using Lyapunov stability analysis," *ISA Transactions*, vol. 98, pp. 292–308, 2020.
- [15] S. Chaturvedi, N. Kumar, and R. Kumar, "A PSO optimized novel PID neural network model for temperature control of jacketed CSTR: design, simulation, and a comparative study," 2023, https://www.researchsquare.com/article/rs-2851632/v1.
- [16] R. Kumar, "A Lyapunov-stability-based context-layered recurrent pi-sigma neural network for the identification of nonlinear systems," *Applied Soft Computing*, vol. 122, Article ID 108836, 2022.
- [17] B. Alyafi, O. Diaz, and R. Marti, "DCGANs for realistic breast mass augmentation in x-ray mammography," in *Medical Imaging 2020: Computer-Aided Diagnosis*, vol. 11314, pp. 473–480, SPIE, Washington, DC, USA, 2020.
- [18] E. Wu, K. Wu, and W. Lotter, "Synthesizing lesions using contextual GANs improves breast cancer classification on mammograms," 2020, https://arxiv.org/abs/2006.00086.
- [19] S. D. Desai, S. Giraddi, N. Verma, P. Gupta, and S. Ramya, "Breast cancer detection using gan for limited labeled dataset," in 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), pp. 34–39, IEEE, Bhimtal, India, September 2020.
- [20] T. Shen, K. Hao, C. Gou, and F.-Y. Wang, "Mass image synthesis in mammogram with contextual information based on GANs," *Computer Methods and Programs in Biomedicine*, vol. 202, Article ID 106019, 2021.
- [21] B. Swiderski, L. Gielata, P. Olszewski, S. Osowski, and M. Kołodziej, "Deep neural system for supporting tumor recognition of mammograms using modified GAN," *Expert Systems with Applications*, vol. 164, Article ID 113968, 2021.

- [22] J. Lee and R. M. Nishikawa, "Identifying women with mammographically-occult breast cancer leveraging GANsimulated mammograms," *IEEE Transactions on Medical Imaging*, vol. 41, no. 1, pp. 225–236, 2022.
- [23] S. Park, K. H. Lee, B. Ko, and N. Kim, "Unsupervised anomaly detection with generative adversarial networks in mammography," *Scientific Reports*, vol. 13, no. 1, p. 2925, 2023.
- [24] U. O. S. Florida, "Digital Database for screening mammography," 2023, http://www.eng.usf.edu/cvprg/mammography/ database.html.
- [25] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, https://arxiv.org/abs/1511.06434.