IPTEK The Journal of Technology and Science, 34(1), 2023 (e/pISSN:2088-2033, 0853-4098)
DOI: 10.12962/j20882033.v34i3.14968 | Received 25 Nov, 2022; Revised 5 Jan, 2023; Accepted 23 Jan, 2021

ORIGINAL RESEARCH

ULTRASOUND IMAGE SYNTHETIC GENERATING USING DEEP CONVOLUTION GENERATIVE ADVERSARIAL NETWORK FOR BREAST CANCER IDENTIFICATION

Dina Zatusiva Haq | Chastine Fatichah*

Dept. of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Correspondence

*Chastine Fatichah, Dept of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. Email: chastine@if.its.ac.id

Present Address

Gedung Teknik Informatika, Jl. Teknik Kimia, Surabaya 60111, Indonesia

Abstract

Breast cancer is the leading cause of death in women worldwide; prevention of possible death from breast cancer can be decreased by early identification ultrasound image analysis by classifying ultrasound images into three classes (Normal, Benign, and Malignant), where the dataset used has imbalanced data. Imbalanced data cause the classification system only to recognize the majority class, so it is necessary to handle imbalanced data. In this study, imbalanced data can be handled by implementing the Deep Convolution Generative Adversarial Network (DCGAN) method as the addition of synthetic images to the training data. The DCGAN method generates synthetic images with feature learning on a Convolutional Neural Network (CNN), making DCGAN more stable than the basic generative adversarial network method. Synthetic and original images were further classified using the CNN GoogleNet method, which performs well in image classification and with reasonable computation cost. Synthetic ultrasound images were generated using a tuning hyperparameter in the DCGAN method to adjust the input size on GoogleNet for imbalanced data handling. From the experiment result, the implementation of DCGAN-GoogleNet has a higher accuracy in handling imbalanced data than conventional augmentation and other previous research, with an accuracy value reaching 91.61%, which is 1% to 4% higher than the accuracy value in the previous method.

KEYWORDS:

Breast Cancer, DCGAN, Imbalanced Data Handling, GoogleNet, Ultrasound Image

1 | INTRODUCTION

Cancer is a disease that ranks first as the deadliest disease in the world[1, 2]. About 20% of teenagers up to 75 years are at risk of being diagnosed with cancer, and about 10% are risking dying from cancer[3] By the number of cancer cases worldwide, breast cancer has the highest number of sufferers and is the leading cause of death in women. The global cancer burden estimates that around 2.36 million people its diagnosed with breast cancer[3, 4]. Increasing the number of breast cancer patients can be overcome by identifying breast cancer. Breast cancer patients are characterized by abnormal cell growth in breast tissue[5] The growth of cancer cells in breast tissue can be caused by factors such as population structure, heredity or genetics, lifestyle, and environment. Various factors create different mortality and survival rates in each region[6] One way to identify the presence of breast cancer is by doing a breast self-examination; in this way, every woman would feel the condition of her breasts based on mass, discharge, swelling, and other possible abnormalities[7] However, this method can lead to overdiagnosis and inaccuracy of identification results. Many cases of breast cancer that attack women worldwide can be prevented by identifying breast cancer at an early stage. Identification of breast cancer can be made by medical image analysis.

Several screening methods can be used to identify cancer cells in the breast, including Magnetic Resource Imaging (MRI)[8], clinical breast examination[9], mammography, and ultrasound[10]. Breast cancer screening based on mammography images is the most often used and studied in previous research studies[11–13]. However, breast cancer screening using mammography images is considered less effective and accurate because of its inability to detect small cancer cells and patients with dense breast tissue[14]. In addition, mammography imaging also requires expensive equipment and equipment maintenance. Ultrasound imagery is a solution to identify breast cancer with a more affordable cost and broader coverage in developing countries[15]. Breast cancer identification can be made by using Computer Aided Diagnosis (CAD).

Identification systems using CAD can be carried out using deep learning methods. Deep learning is a development of machine learning that is more advanced and performs well[16]. Several deep learning methods that are commonly used in disease identification systems include Recurrent Neural Networks (RNN)[1], Deep Belief Network (DBN)[17], and Convolution Neural Network (CNN)[18]. CNN performs better in medical image analysis. The CNN architecture has a convolution layer that can study image features through a filter matrix[19]. This study utilizes the convolution layer to deal with unbalanced data problems using the Generative Adversarial Networks (GAN), called the Deep Convolution Generative Adversarial Networks (DCGAN) method. The DCGAN method generates synthetic images based on the original image features to be added to each class in the dataset so that each class has the same number of images[20]. Besides being used in dealing with imbalanced data problems, in this study, the convolution layer is also used in the feature learning process in the CNN GoogleNet method. GoogleNet is an architecture introduced by Google in 2014 and was ranked first in the 2014 ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition as the architecture with the best performance[21].

2 | PREVIOUS RESEARCHES

Implementation of CAD in the classification of breast cancer ultrasound images in previous studies using the Artificial Network method obtained the highest accuracy of 82.04%[10]. Deep learning is developing the neural network method that is being researched. Convolutional Neural Network (CNN) is a deep learning method that allows the feature learning process and image classification to be carried out in one architecture. Implementation of the CNN method in ultrasound image classification to identify breast cancer in this study[22]. The classification process is carried out with two models. The first model is a pre-trained VGG 16, trained using over 14 million images with 1000 labels. The second model is a fine-tuned of the pre-trained VGG 16 model. The entire experiment shows that a fine-tuned model of medical data can improve accuracy.

In another study, classifying ultrasound images for breast cancer classification used several ResNet architectural models (ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152). The comparison results show that the ResNet-18 architecture has the best accuracy compared to other ResNet architectural models, reaching 88.89%[23].

Previous studies have shown that CNN architectures perform well in ultrasound breast cancer image classification. Comparison of CNN architectures such as VGG-16, VGG-19, ResNet-18, ResNet-50, ResNet 101, AlexNet, and GoogleNet in the Arabic handwritten image classification shows that GoogleNet has a higher accuracy value than other CNN architectures. The classification system generated by the GoogleNet architecture achieves the highest accuracy of 95.5%[24]. Based on previous research, the GoogleNet architecture has a good ability in image classification.

The ultrasound image dataset of breast cancer used in this study is an imbalanced data that can affect the accuracy of results. Class prediction results on imbalanced data tend to be predicted as the majority class or the most data with a significant difference in the amount of data[25]. Handling the imbalance in image data can be done by augmenting the image in the minority class so that

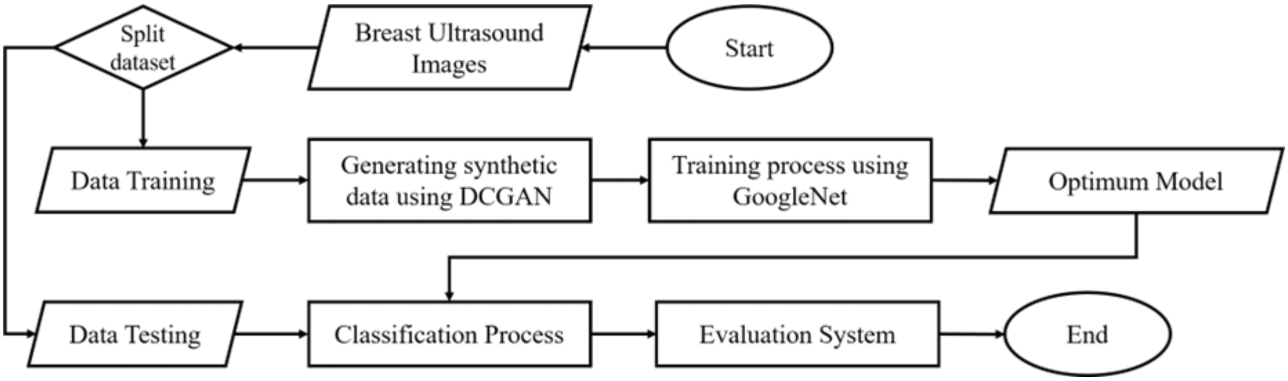


FIGURE 1 The research flowchart.

each class has a balanced amount of data. Research by Malygina et al.[26] on cases of imbalanced Chest X-Ray data to predict pathologies uses data augmentation by adding a data set with synthetic data using the Generative Adversarial Network (GAN) method. This research shows the addition of synthetic images to handle imbalanced data problems and increase the accuracy by 2% in the classification process. In generator artificial neural network in GAN architecture, the probability density distribution studied does not show clear image results, and the generator tends to be unstable[27]. The development of the GAN method to overcome its weaknesses is the Deep Convolution Generative Adversarial Networks (DCGAN) method[28].

Research conducted by Rejusha and KS[29] shows that DCGAN can enhance the accuracy of the classification process compared to other augmentation techniques, such as Affine Image Transformations and Elastic transformations. In this study, the handling of imbalanced data with DCGAN was carried out on the classification of breast cancer ultrasound images using GoogleNet. This method can give a better breast cancer identification system in accuracy.

Based on the exposure of previous studies, in this study, the handling of imbalanced data with DCGAN was carried out to classify breast cancer ultrasound images using GoogleNet. This research is expected to produce a breast cancer identification system with reasonable accuracy.

3 | MATERIAL AND METHOD

The research on breast identification system based on breast ultrasound image classification using DCGAN and hybrid GoogleNet methods generally consist of three stages. The research design in this study contains the steps of work carried out following the research framework in Fig. 1

3.1 | Dataset Description

The dataset used to study breast cancer identification systems based on ultrasound images was obtained from[30]. Data was taken in 2018 from a sample of 600 women aged 25 to 75 years. The data consists of 780 images divided into three classes: normal, benign, and malignant breast ultrasound. Normal class breast ultrasound images totaled 487, in this class showing healthy breasts. In comparison, the benign and malignant classes indicate the presence of early and advanced breast cancer with the amount of data in each class of 210 and 133. The average size of the image is 500×500, and in the first step, each image is resized to 224×224, appropriate as the input size in GoogleNet. An example of an ultrasound image is shown in Fig 2 .

3.2 | Splitting Data

K-fold cross-validation is a method used to divide training and testing data with a proportional percentage of distribution[31]. Compared to other methods, the k-fold cross-validation method produces a model with less bias[32]. In k-fold cross-validation, the dataset is randomly separated into k subsets of the same size, and the method is repeated k times. Each iteration of one data set is used for testing data and the rest for training data[33]. An illustration of the distribution of training data and testing data with k=5 is shown in Fig. 3

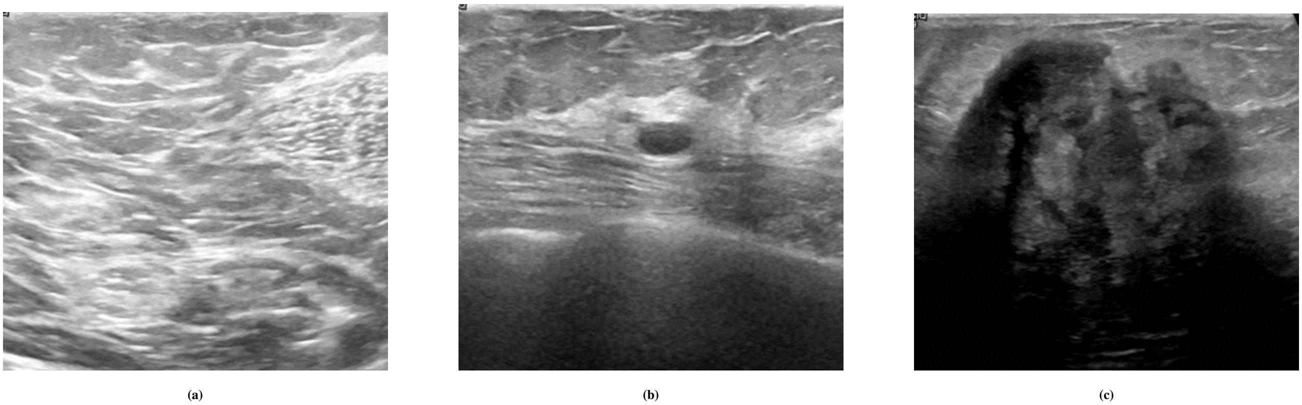


FIGURE 2 The example of ultrasound images: (a) Breast ultrasound normal image, (b) Breast ultrasound benign image, (c) Citra breast.

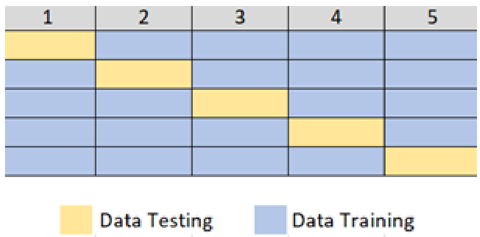


FIGURE 3 The illustration of the distribution of training data and testing data with k=5.

3.3 | Deep Convolution Generative Adversarial Networks (DCGAN)

Generative Adversarial Networks (GAN) were introduced in 2014 by Goodfellow et al.[34]. GAN is a deep learning model that can generalize synthetic images similar to original images[35]. GAN has two neural networks, a generator, and a discriminator; it is optimized simultaneously but with the opposite task. The generator artificial neural network synthesizes the image as closely as possible to the original image so that the features in the synthesized image are still recognized as the original image. In contrast, the discriminator artificial neural network distinguishes the image from the original image to increase the dataset's variance[36]. However, the probability density distribution studied in the generator artificial neural network does not show clear image results, and the generator tends to be unstable[27]. The development of the GAN method to overcome its weaknesses is the Deep Convolution Generative Adversarial Networks (DCGAN) method[28].

3.4 | GoogleNet

GoogleNet is a CNN architecture introduced by Google in 2014 and won first place in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014 competition as the best-performing architecture. The error value obtained by GoogleNet in the classification of image data reaches 6.7%[37] The GoogleNet architectural learning model can achieve high accuracy by deepening layers to improve neural network performance[38] GoogleNet has 144 layers with the image size at the input layer 224×224×3[39].

The advantage of GoogleNet is in the inception modules that CNN does not have. Inception modules consist of several small convolutions that aim to reduce the number of parameters without reducing network performance[40]. The output in the previous layer is processed in 4 layers of 1×1 convolution layer, one layer of 3×3 convolution layer, one layer of 5×5 convolution layer, and 3×3 pooling, which is added together with the arrangement as shown in Fig. 4 [41].

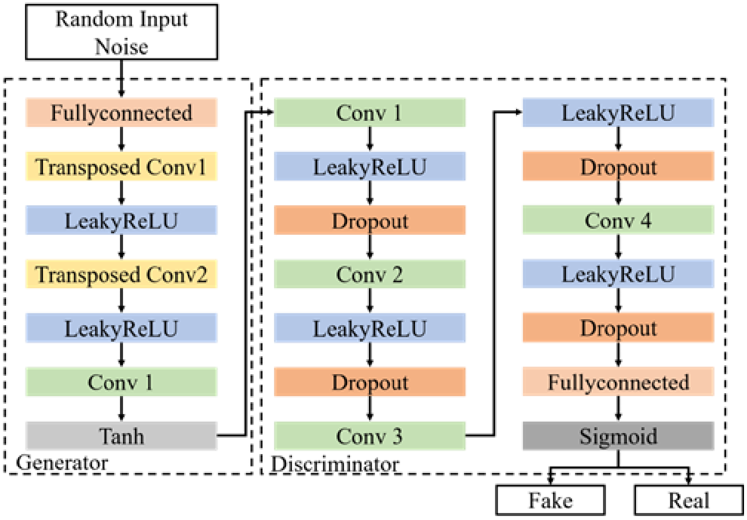


FIGURE 4 The illustration of data splitting using k-fold cross-validation ultrasound malignant.

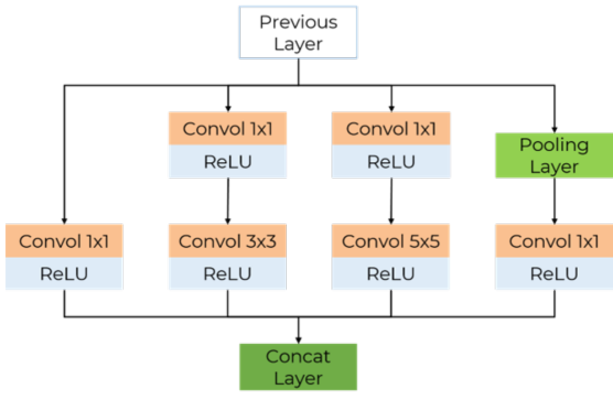


FIGURE 5 The inception modules in GoogleNet Architecture.

TABLE 1 The confusion matrix.

		Predicted		
		Normal	Benign	Malignant
Actual	Normal	TP	FP	FP
	Benign	FN	TN	TN
	Malignant	FN	TN	TN

3.5 | Evaluation System

The confusion matrix is the evaluation system used to find out how well the classification method performs. The confusion matrix is nxn, with n being the number of classes. The confusion matrix measures the performance of the classification system based on the amount of data classified incorrectly using four parameters, namely True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN)[42]. In the medical data classification, FP is defined as the amount of normal data that predicts benign or malignant. In contrast, FN is the data in the breast cancer class but is predicted as normal[43]. The confusion matrix table is shown in Table 1.

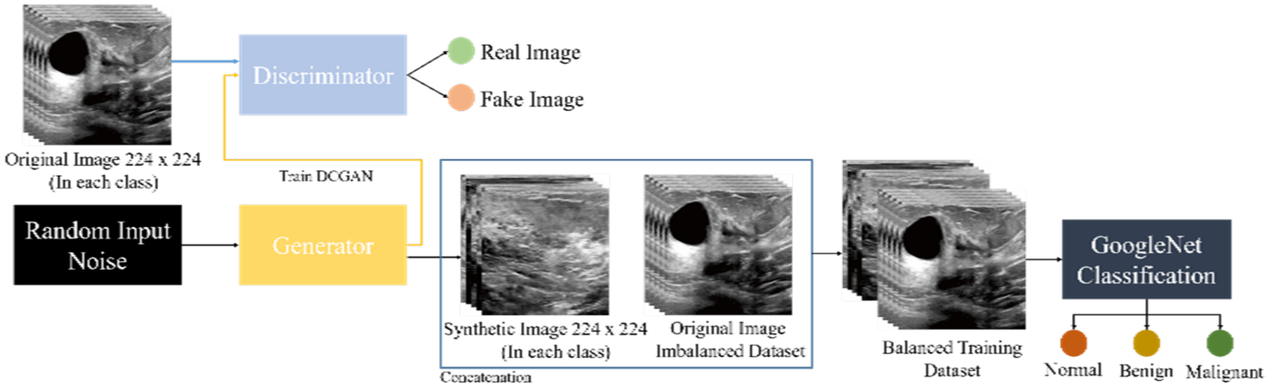


FIGURE 6 The process of adding variance of ultrasound image data to training data.

The parameters in the confusion matrix are used to calculate accuracy (AC), sensitivity (SE), and specificity (SP) with Equations 1-3 to find out how well the classification system performs. The higher the value of AC, SE, and SP, the better the performance of the classification system.

$$Accuracy = \frac{\frac{TP+TN}{TP+FN+FP+TN}}{n} \tag{1}$$

$$Specificity = \frac{\frac{TN}{TN+FP}}{n} \tag{2}$$

$$Sensitivity = \frac{\frac{TP}{TP+FN}}{n} \tag{3}$$

The SE or true-positive rate is a probability value of a normal class ultrasound image but identified as a breast cancer image (benign and malignant). SP or false-positive rate is a probability value of a breast cancer image but identified as a normal class image.

4 | RESULTS AND DISCUSSION

The stage of this research begins with pre-processing the data. Data pre-processing consists of two stages: split data and data augmentation. The distribution of training and testing data is carried out with 5-fold cross-validation. The number of training data is 625, with the number of benign, malignant, and normal classes as many as 350, 168, and 107, respectively. In contrast, the testing data is 155, with details of 87 benign, 42 malignant, and 26 normal data. The number between classes is not balanced with a difference of 243 data. The imbalanced dataset affects the classification results, which would be recognized as the class with the largest data. So it is necessary to add variance of synthetic image data by implementing the Deep Convolution Generative Adversarial Network (DCGAN) method.

The DCGAN method utilizes convolution layers to study image features to create a synthetic image with the same characteristics as the original image. The process of adding variance of ultrasound image data to training data is illustrated in Fig. 6 .

Based on Fig. 6 , the generator and discriminator work simultaneously with random noise input on a generator and original image input on the discriminator. The original input image on the discriminator is taken randomly to produce a synthetic image with the same class. Furthermore, the original and synthetic images are combined into training data so that the training data has the same number in each class.

The size of the convolution layer on the DCGAN architecture is suitable for the size of the input on the GoogleNet architecture. The adjusted DCGAN architecture would produce a synthetic image of 224×224 size. The resulting image, according to the

TABLE 2 The DCGAN architecture with output size 224 x 224.

Layers Name	**Type**	**Learnable Properties**
	Generator	
Fully connected	Fully Connected Layer	Weight 401408×100 Bias 401408×1
TransConv 1	Transposed Convolution Layer	Weight 3×3×128×128 Bias 128×1 Stride 2
TransConv 2		Weight 3×3×64×128 Bias 64×1 Stride 2
Conv 1	Convolution Layer	Weight 3×3×64 Bias 1×1 Stride 1
	Discriminator	
Conv 1	Convolution Layer	Weight 3×3×1×32 Bias 32×1 Stride 2
Conv 2		Weight 3×3×32×64 Bias 64×1 Stride 2
Conv 3		Weight 3×3×32×128 Bias 128×1 Stride 2
Conv 4		Weight 3×3×32×256 Bias 128×1 Stride 2
Fully connected	Fully Connected Layer	Weight 1×4096 Bias 1×1

input size on GoogleNet can reduce the possibility of losing the essence of image features due to a significant resizing process. The classification process identifies breast ultrasound into three stages, normal, benign, and malignant. A description of the size of the convolution layer on the DCGAN architecture is shown in Table 2 .

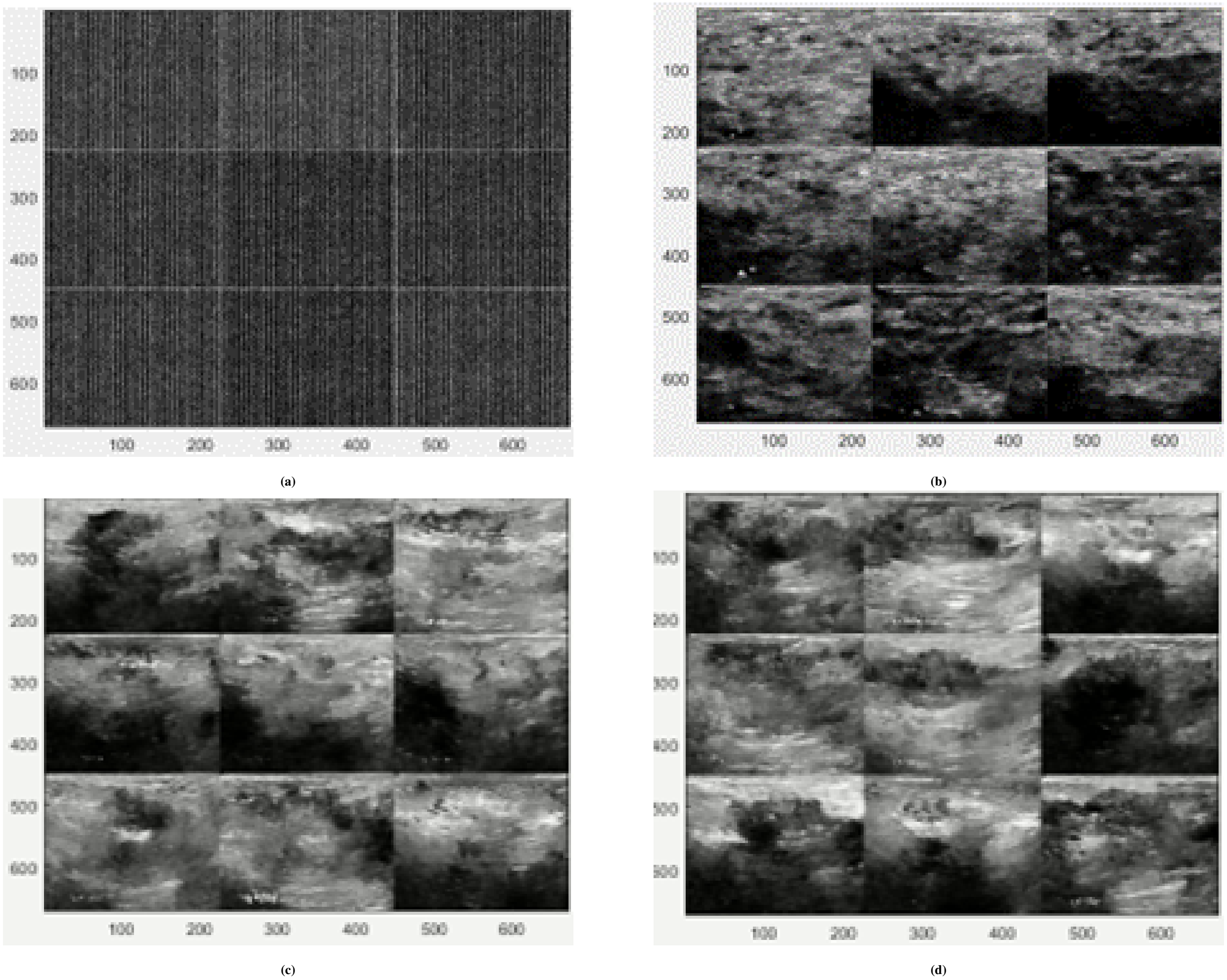
The number of epochs significantly influences the performance of the DCGAN method. The greater the number of epoch values, the better the synthetic image results, but it also takes a long time. An example of synthetic image generalization at epochs 0, 1000, 1500, and 2000 is shown in Fig. 7 .

Based on Figure 6, the generalization of synthetic images using the DCGAN method produces nine synthetic images. In Figure 6 (a), in the first epoch, the image still does not look like an ultrasound image because the image still contains random noise. In Figure 6 (b), using epoch 1000, the image is slightly closer to the original ultrasound image but is still less clear. While in Fig. 7 (c) and (d), the image resembles the original ultrasound image. However, in Fig. 6 (d), the ultrasound image is more stable than the image with 1500 epochs. Each class's synthetic ultrasound image samples are shown in Fig. 8 , Fig. 9 , and Fig. 10 .

Based on Fig. 8 , Fig. 9 , and Fig. 10 , the synthetic image visually resembles the original image. As in the benign class, the synthetic and original images have the same characteristics, like a small black asymmetrical circle. In the malignant class, the black asymmetrical shape looks bigger than in the benign class. Furthermore, synthetic images are added to the training data so that the image in each class is the same. In the normal class, 50 synthetic data were added, while in the benign and malignant classes, 232 and 293 were added. So in total, there are 400 data in each class in the training process. The comparison of the amount of training data in each class after adding synthetic images is shown in Fig. 11 .

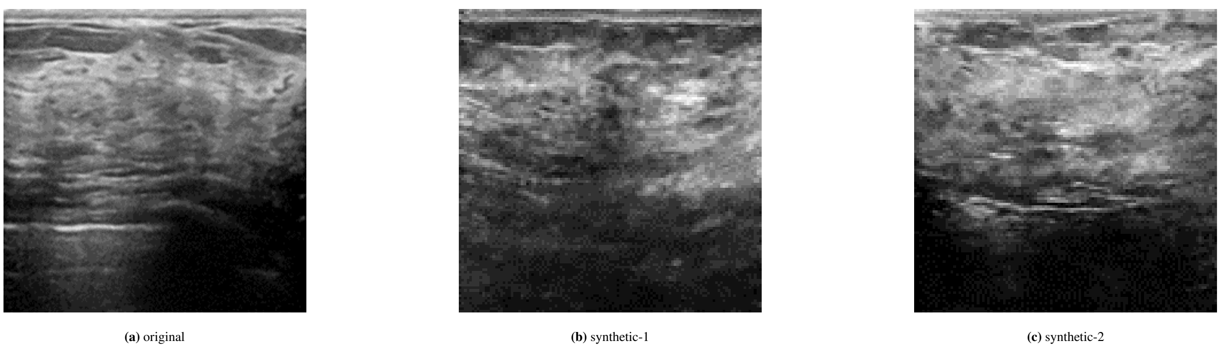
Based on the graph in Fig. 7 , the training data has a balanced dataset after adding synthetic images to each class. The next process is image classification using CNN GoogleNet architecture.

The image classification resulting from the addition of synthetic images is compared with the addition of images using conventional augmentation methods. In experiment 6, the conventional augmentation method used as a comparison is by randomly rotating the image at a special angle (30, 45, 60, and 90 degrees) so that the number of images is the same as the amount of training data that has been added with synthetic images. In experiment 7, the synthetic images were generated by an unsuitable DCGAN method. The unsuitable DCGAN method produces synthetic images more minor than the input size in GoogleNet,



(a) (b) (c) (d)

FIGURE 7 Example of synthetic image generalization at epochs (a) 1, (b) 1000, (c) 1500, and (d) 2000.



(a) original **(b)** synthetic-1 **(c)** synthetic-2

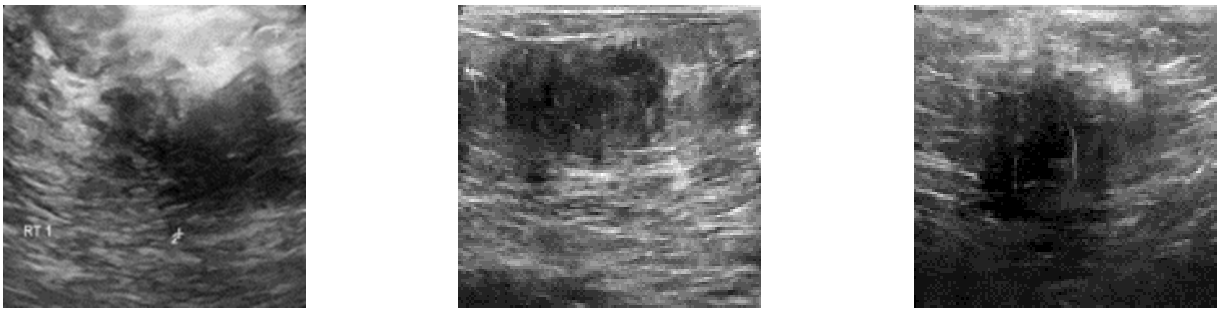
FIGURE 8 The synthetic ultrasound image samples for 'Normal' class.

which is 28×28. Comparison of classification results based on evaluation using accuracy (AC), sensitivity (SE), and specificity (SP) in Equation 1-3. The system evaluation of each experiment is shown in Table 3 .

Based on Table 3 , implementing the DCGAN method to handle imbalanced data and GoogleNet as an image classification method got better accuracy than the Geometry Augmentation Transformation-GoogleNet and previous studies with an accuracy of 91.61%. In research[44], the GAN-CNN method obtained a fairly good accuracy reaching 90.41%, but this experiment only

(a) original **(b)** synthetic-1 **(c)** synthetic-2

FIGURE 9 The synthetic ultrasound image samples for 'Benign' class.



(a) original **(b)** synthetic-1 **(c)** synthetic-2

FIGURE 10 The synthetic ultrasound image samples for 'Malignant' class.

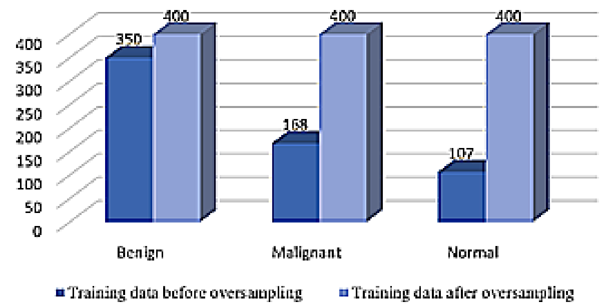


FIGURE 11 The comparison number of data in the training dataset before and after imbalanced handling.

tested using two classes of ultrasound data. Research by Khanna et al.[47] extracted ultrasound image features using the feature learning layer on the ResNet50 architecture. The feature extraction vector array is then carried out by the feature selection and classification process using the Support Vector Machine (SVM) method. This study obtained a fairly good accuracy of 84.90%.

The performance of the classification system is also determined by SE and SP values. In medical image classification, The SE value is highly considered. When the SE value is high, it reduces the possibility of breast cancer not being identified. A high SE value indicates that the system can recognize ultrasound images well, so most misclassifications are normal data but are identified as breast cancer. In classifying medical images, errors like this are better than patients diagnosed with breast cancer but not identified. The SP value indicates that the system can recognize normal class data well, indicating some benign or malignant data as a normal class. Analyzing SE and SP values determines the best classification method for SE values. The highest SE value was in experiment 8, with an SE value of 94.20%. The visualization of the accuracy results is shown in Fig. 12 .

TABLE 3 The comparison of the classification evaluation system.

Method	Dataset	AC(%)	SE(%)	SP(%)
GAN-CNN [44]	University of Malaya Medical Centre (2 classes)	90.41	87.94	85.86
ResNet [45]	BUSI Dataset	83.00	-	-
BVA Net [46]		84.30	75.80	88.30
ResNet50-SVM with feature selection [47]		84.90	-	-
GoogleNet		78.06	80.84	78.82
Transformation Geometry augmentation-GoogleNet		87.74	84.94	89.60
Unsuitable DCGAN-GoogleNet		81.94	75.38	83.46
DCGAN-GoogleNet		91.61	94.20	89.74

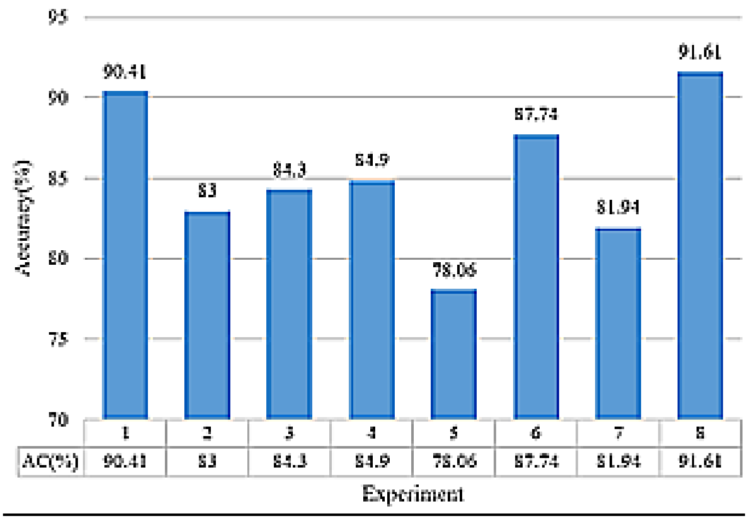


FIGURE 12 The visualization of the accuracy results.

TABLE 4 The confusion matrix of the best model.

		Predicted Normal	Predicted Benign	Predicted Malignant
Actual	Normal	76	2	0
Actual	Benign	6	40	0
Actual	Malignant	4	0	26

Experiment 8 has a higher accuracy than Experiment 6. Based on all experiments, generating a synthetic image is a better way to add variants of image data than the conventional augmentation method. Synthetic images by the DCGAN method reduce the possibility of missing image features due to the augmentation process by creating new images based on the original image features.

Based on Fig. 12 , it can be seen that in experiment 8 using DCGAN-GoogleNet, the highest accuracy was achieved, 91.61%, followed by the first and sixth experiments with an accuracy of 90.41% and 87.74%, respectively. The results of the confusion matrix experiment 8 are shown in Table 4

Based on Table 4, 76 benign data, 40 malignant data, and 26 normal data were classified correctly. 13 data are classified as not in the actual class. The most errors are in the malignant class, where seven classes should be included in the benign class but are classified as malignant. It shows that the similarity between malignant and benign data is higher than in other classes. The high accuracy results in Experiment 8 prove that the synthetic image generated using the DCGAN method can be recognized as a genuine ultrasound image.

All experiments show that GoogleNet has a good performance in image classification. In the GoogleNet architecture, feature learning and classification processes occur. The feature learning process is an image feature extraction process through learning the image's shape, edge, and color features. In general, CNN performs well in image classification but has drawbacks in the length of training time and high computational complexity [48]. Research by [49] combines the feature learning process on the CNN architecture to reduce training time in the CNN architecture with the Extreme Learning Machine (ELM) classification method. The combination of the CNN-ELM method shows better results than the conventional CNN method, with a difference in the accuracy of 2 to 4% and a training time range of 4000 seconds, so the combination of CNN-ELM can be a solution to overcome the problem of long training times in the CNN architecture.

5 | CONCLUSION

The DCGAN method can handle the imbalanced data problem well based on the trial results. The 2000 number of epochs in the DCGAN method can generalize the synthetic ultrasound image well. The synthetic image has a visual resemblance to the original image. As in the benign class, the synthetic and original images have the same characteristics, like a small black asymmetrical circle. In the malignant class, the black asymmetrical shape looks bigger than in the benign class. The classification results using the GoogleNet method show better accuracy than the Geometry Transformation augmentation-GoogleNet and previous studies with an accuracy of 91.61%.

CREDIT

Dina Zatusiva Haq: Conceptualization, Methodology, Software, Writing. **Chastine Fatichah:** Supervision, Validation

References

1. Asyhar AH, Foeady AZ, Thohir M, Arifin AZ, Haq DZ, Novitasari DCR. Implementation LSTM Algorithm for Cervical Cancer using Colposcopy Data. In: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) IEEE; 2020. p. 485–489.

2. Pullaiah NRP, Venkatasekhar D, Venkatramana P, Sudhakar B. Detection of Breast Cancer on Magnetic Resonance Imaging Using Hybrid Feature Extraction and Deep Neural Network Techniques. International Journal of Intelligent Engineering and Systems 2020;13(6):229–240.

3. Ferlay J, Colombet M, Soerjomataram I, Parkin DM, Piñeros M, Znaor A, et al. Cancer statistics for the year 2020: An overview. International Journal of Cancer 2021;149(4):778–789.

4. Wilkinson L, Gathani T. Understanding breast cancer as a global health concern. The British Journal of Radiology 2022;95(1130):20211033.

5. Vijayarajeswari R, Parthasarathy P, Vivekanandan S, Basha AA. Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. Measurement 2019;146:800–805.

6. Momenimovahed Z, Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. Breast Cancer: Targets and Therapy 2019;11:151.

7. Dagne AH, Ayele AD, Assefa EM. Assessment of breast self-examination practice and associated factors among female workers in Debre Tabor Town public health facilities, North West Ethiopia, 2018: Cross-sectional study. PloS one 2019;14(8):e0221356.

8. Mann RM, Kuhl CK, Moy L. Contrast-enhanced MRI for breast cancer screening. Journal of Magnetic Resonance Imaging 2019;50(2):377–390.

9. Bhan AD, Jayaram J. Screening, Self-Examination and Awareness in Breast Cancer. In: Breast Cancer Springer; 2022.p. 587–600.

10. Mohammed MA, Al-Khateeb B, Rashid AN, Ibrahim DA, Abd Ghani MK, Mostafa SA. Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. Computers & Electrical Engineering 2018;70:871–882.

11. Kavitha T, Mathai PP, Karthikeyan C, Ashok M, Kohar R, Avanija J, et al. Deep learning based capsule neural network model for breast cancer diagnosis using mammogram images. Interdisciplinary Sciences: Computational Life Sciences 2022;14(1):113–129.

12. Milosevic M, Jankovic D, Milenkovic A, Stojanov D. Early diagnosis and detection of breast cancer. Technology and Health Care 2018;26(4):729–759.

13. Htay TT, Maung SS. Early stage breast cancer detection system using glcm feature extraction and k-nearest neighbor (k-NN) on mammography image. In: 2018 18th International Symposium on Communications and Information Technologies (ISCIT) IEEE; 2018. p. 171–175.

14. Li J, Guan X, Fan Z, Ching LM, Li Y, Wang X, et al. Non-invasive biomarkers for early detection of breast cancer. Cancers 2020;12(10):2767.

15. Sood R, Rositch AF, Shakoor D, Ambinder E, Pool KL, Pollack E, et al. Ultrasound for breast cancer detection globally: a systematic review and meta-analysis. Journal of global oncology 2019;5:1–17.

16. Liu H, Lang B. Machine learning and deep learning methods for intrusion detection systems: A survey. applied sciences 2019;9(20):4396.

17. Novitasari, Foeady AZ, Thohir M, Arifin AZ, Niam K, Asyhar AH. Automatic Approach for Cervical Cancer Detection Based on Deep Belief Network (DBN) Using Colposcopy Data. In: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) IEEE; 2020. p. 415–420.

18. Ahmed N, Yigit A, Isik Z, Alpkocak A. Identification of leukemia subtypes from microscopic images using convolutional neural network. Diagnostics 2019;9(3).

19. Arya S, Singh R. A Comparative Study of CNN and AlexNet for Detection of Disease in Potato and Mango leaf. In: 2019 International conference on issues and challenges in intelligent computing techniques (ICICT), vol. 1 IEEE; 2019. p. 1–6.

20. Dewi C, Chen RC, Liu YT, Tai SK. Synthetic Data generation using DCGAN for improved traffic sign recognition. Neural Computing and Applications 2022;34(24):21465–21480.

21. Chelghoum R, Ikhlef A, Hameurlaine A, Jacquir S. Transfer learning using convolutional neural network architectures for brain tumor classification from MRI images. In: IFIP International Conference on Artificial Intelligence Applications and Innovations Springer; 2020. p. 189–200.

22. Hijab A, Rushdi MA, Gomaa MM, Eldeib A. Breast cancer classification in ultrasound images using transfer learning. In: 2019 Fifth International Conference on Advances in Biomedical Engineering (ICABME) IEEE; 2019. p. 1–4.

23. Das A, Rana S. Exploring Residual Networks for Breast Cancer Detection from Ultrasound Images. In: 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) IEEE; 2021. p. 1–6.

24. Almisreb AA, Turaev S, Saleh MA, Al Junid SAM. Arabic Handwriting Classification using Deep Transfer Learning Techniques. Pertanika Journal of Science & Technology 2022;30(1).

25. Thabtah F, Hammoud S, Kamalov F, Gonsalves A. Data imbalance in classification: Experimental evaluation. Information Sciences 2020;513:429–441.

26. Malygina T, Ericheva E, Drokin I. Data augmentation with GAN: Improving chest X-ray pathologies prediction on class-imbalanced cases. In: International conference on analysis of images, social networks and texts Springer; 2019. p. 321–334.

27. Komatsu R, Gonsalves T. Conditional DCGAN's challenge: generating handwritten character digit, alphabet and katakana. In: Proceedings of the Annual Conference of JSAI 33rd (2019) The Japanese Society for Artificial Intelligence; 2019. p. 3B3E204–3B3E204.

28. Öcal A, Özbakır L. Supervised deep convolutional generative adversarial networks. Neurocomputing 2021;449:389–398.

29. Rejusha TR, KS VK. Artificial MRI Image Generation using Deep Convolutional GAN and its Comparison with other Augmentation Methods. In: 2021 International Conference on Communication, Control and Information Sciences (ICCISc), vol. 1 IEEE; 2021. p. 1–6.

30. Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. Data in brief 2020;28:104863.

31. Sanjaya S, Pura ML, Gusti SK, Yanto F, Syafria F. K-Nearest Neighbor for Classification of Tomato Maturity Level Based on Hue, Saturation, and Value Colors. Indonesian Journal of Artificial Intelligence and Data Mining 2019;2(2):101.

32. Rodriguez JD, Perez A, Lozano JA. Sensitivity analysis of k-fold cross validation in prediction error estimation. IEEE transactions on pattern analysis and machine intelligence 2009;32(3):569–575.

33. Ziasabounchi N, Askerzade I. ANFIS based classification model for heart disease prediction. International Journal of Electrical & Computer Sciences IJECS-IJENS 2014;14(02):7–12.

34. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Communications of the ACM 2020;63(11):139–144.

35. Sorin V, Barash Y, Konen E, Klang E. Creating artificial images for radiology applications using generative adversarial networks (GANs)–a systematic review. Academic radiology 2020;27(8):1175–1185.

36. Wolterink JM, Mukhopadhyay A, Leiner T, Vogl TJ, Bucher AM, Išgum I. Generative Adversarial networks: a primer for radiologists. Radiographics 2021;41(3):840–857.

37. Seo Y, Shin Ks. Hierarchical convolutional neural networks for fashion image classification. Expert Systems with Applications 2019;116:328–339.

38. Kim JH, Seo SY, Song CG, Kim KS. Assessment of electrocardiogram rhythms by GoogLeNet deep neural network architecture. Journal of healthcare engineering 2019;2019.

39. Zhang X, Pan W, Bontozoglou C, Chirikhina E, Chen D, Xiao P. Skin capacitive imaging analysis using deep learning GoogLeNet. In: Science and Information Conference Springer; 2020. p. 395–404.

40. Jasitha P, DIleep MR, DIvya M. Venation Based Plant Leaves Classification Using GoogLeNet and VGG. 2019 4th IEEE International Conference on Recent Trends on Electronics, Information, Communication and Technology, RTEICT 2019 - Proceedings 2019;p. 715–719.

41. Zmudzinski L. Deep Learning Guinea Pig Image Classification Using Nvidia DIGITS and GoogLeNet. In: CS& P; 2018. p. 5–15.

42. Koklu M, Ozkan IA. Multiclass classification of dry beans using computer vision and machine learning techniques. Computers and Electronics in Agriculture 2020;174:105507.

43. Ibrahim AU, Ozsoz M, Serte S, Al-Turjman F, Yakoi PS. Pneumonia classification using deep learning from chest X-ray images during COVID-19. Cognitive Computation 2021;p. 1–13.

44. Pang T, Wong JHD, Ng WL, Chan CS. Semi-supervised GAN-based Radiomics Model for Data Augmentation in Breast Ultrasound Mass Classification. Computer Methods and Programs in Biomedicine 2021;203:106018. https://www.sciencedirect.com/science/article/pii/S0169260721000936.

45. Gheflati B, Rivaz H. Vision Transformers for Classification of Breast Ultrasound Images. In: 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); 2022. p. 480–483.

46. Xing J, Chen C, Lu Q, Cai X, Yu A, Xu Y, et al. Using bi-rads stratifications as auxiliary information for breast masses classification in ultrasound images. IEEE Journal of Biomedical and Health Informatics 2020;25(6):2058–2070.

47. Khanna P, Sahu M, Singh BK. Improving the classification performance of breast ultrasound image using deep learning and optimization algorithm. In: 2021 IEEE International Conference on Technology, Research, and Innovation for Betterment of Society (TRIBES); 2021. p. 1–6.

48. Xu K, Feng D, Mi H. Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image. Molecules 2017;22(12):2054.

49. Khellal A, Ma H, Fei Q. Convolutional neural network based on extreme learning machine for maritime ships recognition in infrared images. Sensors 2018;18(5):1490.

How to cite this article: Haq D.Z., Fatichah C. (2023), Ultrasound Image Synthetic Generating Using Deep Convolution Generative Adversarial Network For Breast Cancer Identification , *IPTEK The Journal of Technology and Science*, *34(1):12-25*.