BỘ GIÁO DỤC VÀ ĐÀO TẠO

TRƯỜNG ĐẠI HỌC CÔNG NGHỆ KỸ THUẬT

THÀNH PHỐ HỒ CHÍ MINH

NGUYỄN THANH TÂM

PHÁT TRIỂN HỆ THỐNG HỖ TRỢ CHẨN ĐOÁN BỆNH UNG THƯ VÚ ỨNG DỤNG TRÍ TUỆ NHÂN TẠO

CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ

Tp. Hồ Chí Minh, tháng 1 năm 2026

DANH MỤC CÔNG TRÌNH ĐÃ CÔNG BỐ

Stt	Tên bài báo	Tác giả	Tên tạp chí/kỷ yếu, số/tập, số ISSN/ISBN	Thời gian xuất bản	Link/Doi	Ghi chú
I.	**Tạp chí khoa học trong danh mục WoS (SCIE, SSCI, AHCI, ESCI)**					
1	Largest roi segmentation for breast cancer classification using a VGG16 deep learning network	Thanh-Tam Nguyen, Thanh-Hai Nguyen, Ba-Viet Ngo, Thanh-Nghia Nguyen	Advances in Electrical and Electronic Engineering, 31 Dec 2024, Vol. 22, Issue 4 ISSN: 18043119	2024	10.15598/ aeee.v22i4. 240303	ESCI Q3 SJR= 0.229
II	**Tạp chí khoa học trong danh mục Scopus**					
2	A Robust Approach for Breast Cancer Classification from DICOM Images	Thanh-Nghia Nguyen, Thanh-Tam Nguyen, Thanh-Hai Nguyen, Ba-Viet Ngo	Engineering, Technology and Applied Science Research, 04 Jun 2025, Vol. 15, Issue 3, pages 23499 – 23505 ISSN: 22414487	2025	10.48084/ etasr.10931	Q2 SJR= 0.332
3	Segmentation and Color ROI Extraction from Breast Imaging Datasets for Cancer Classification	Thanh-Tam Nguyen, Thanh-Hai Nguyen	Engineering, Technology and Applied Science Research, 2025, ISSN: 22414487	2025	10.48084/ etasr.12067	Q2 SJR= 0.332
III	**Tạp chí khoa học trong nước**					
4	Design of A Telemedicine System for Classification	Thanh-Tam Nguyen, Nguyen Thanh	Journal of Technical Education Science	12/09/2025	10.54644/ jte.2025.1969	Đã chấp nhận đăng

	of Breast Cancer Images	Hai, Tin-Trung Nguyen	P-ISSN: 2615-9740			
IV	**Kỷ yếu hội nghị khoa học quốc tế có phản biện và có ISBN**					
5	A GLCM Algorithm for Optimal Features of Mammographic Images for Detection of Breast Cancer	Thanh-Tam Nguyen; Thanh-Hai Nguyen; Ba-Viet Ngo	2021 International Conference on System Science and Engineering (ICSSE), 2021, pp. 295–299 ISBN:978-1-6654-4848-2	2021	10.1109/ ICSSE52999. 2021.9538426	

Research Article

LARGEST ROI SEGMENTATION FOR BREAST CANCER CLASSIFICATION USING A VGG16 DEEP LEARNING NETWORK

Thanh-Tam NGUYEN[1], *Thanh-Hai NGUYEN*[1,*], *Ba-Viet NGO*[1], *Thanh-Nghia NGUYEN*[1]

[1]Department of Industrial Electronics and Biomedical Engineering, Faculty of Electrical – Electronics Engineering, HCMC University of Technology and Education, Ho Chi Minh City, Vietnam

tamnt.ncs@hcmute.edu.vn, nthai@hcmute.edu.vn, vietnb@hcmute.edu.vn, nghiant@hcmute.edu.vn

*Corresponding author: Thanh-Hai NGUYEN; nthai@hcmute.edu.vn

DOI: 10.15598/aeee.v22i4.240303

Article history: Received Mar 17, 2024; Revised Jun 18, 2024; Accepted Aug 08, 2024; Published Dec 31, 2024.

Abstract. *The exact evaluation of breast cancer images for patients is very important, because they can be early treated for lasting their life. This article proposes a classification system for finding breast cancer images, in which each breast lesion image is segmented to produce a largest Region of Interest (ROI) and a VGG16 deep learning network is applied for classification. An Otsu threshold is utilized on two datasets from two sources of CBIS-DDSM and MIAS to create largest ROI with main features. For the classification with high performance, two datasets of the breast lesions were augmented by rotating, flipping, and brightness variation. This article was proposed an algorithm with processing images sets before classification using VGG16. In particular, the results of the largest ROI datasets for four types of breast lesions were represented through segmentation, normalization and enhancement. In addition, the results of classifying four types of breast lesions (BC, BM, MC, MM) were evaluated using confusion matrix, with the high accuracy of around 95%. Another evaluation was that these image sets without ROI/with ROI parts/With the largest ROI only using the Otsu segmentation were compared and the highest accuracy was of the image sets with the largest ROI. The results with the high accuracy demonstrated to illustrate the effectiveness of the proposed method. It means that this method can be developed to classify many stages of breast cancers during diagnosis and treatment.*

Keywords

Breast lesion classification, Data Augmentation, VGG16 deep learning network, Largest ROI, Two datasets of CBIS-DDSM and MIAS.

1. Introduction

Breast cancer in humans is one of the most dangerous diseases and also it is very difficult to treat [1]. The cause of breast cancer is still unclear, so it is very necessary to recognize soon for treatment. However, many researches have suggested that mutations of genes in the BRCA1 or BRCA2 are the main causes [2] such as environmental pollution, radiation, chemicals in food, smoking habits and other issues, which can increase the risk for having breast cancer treatment.

With many different cancers, breast cancer often appears in women, the majority of women are over 40 years old [3]. In particular, 2.3 million women were diagnosed with breast cancer and 685,000 died globally in 2020. About the end of 2020, there were 7.8 million women alive, who had been diagnosed with breast cancer in the past five years [4]. However, if cancers with early diagnosis can be detected, human can have a very high chance of treatment [5]. In particular, if the cancers are early detected in the first stage, the treatment rate is 80%, while second stage is about 60% [6].

An Otsu segmentation algorithm is often employed for determining thresholds to be able to ROI [7]. In particular, the Otsu segmentation is to determine gray level, where a gray level image is calculated to divide pixels into two groups: background pixels, and object pixels. In addition, breast images with many different parts on them are difficult for classifying high performance. Therefore, creating ROI from the breast image using segmentation methods is very necessary [8, 9, 10]. In this research, the largest ROI with features is detected before training VGG16. It is obvious that the obtained results can be better compared to previous works with the same datasets.

Image enhancement using histogram equalization method will make the image clearer and extract features [11]. It means that a histogram equalization technique is often used for enhancing contrast in a variety of different image types such as medical and satellite images. There are many methods of histogram equalization and essentially they are divided into two types of global and local histogram equalization. In particular, the algorithm proposed, in [11], used Gaussian Mixture Model for the model of the gray level distribution of the image. Furthermore, the intersections of the Gaussian Model were used to model the dynamic range of the images into input gray level intervals. Another research is that based on histogram equalization and superpixel segmentation, the method was introduced to find whether an image with grayscale. Thus, this paper can infer whether an image is dark, bright, or a small dynamic range [12].

In recent years, there has been a lot of researches on Artificial Intelligence (AI) in the medical field, particularly applying AI technology in identifying and classifying diseases such as heart [13, 14], human skin [15], brain cancer [16], breast cancer [17] and others. This article proposed an automatic segmentation method based on CNN with a 3x3 kernel. With this method of using a small kernel, it allows the design of a deep learning network architecture and also limits overfitting. The proposed method was applied on the CNN network for the MRI image set of the Brain Tumor Segmentation Challenge 2013 database (BRATS 2013) [16]. In these reseahces, AI has been applied for recognition of breast cancer and evaluating the level of cancers based on mammograms. In addition, using AI to classify and identify cancers, the cancers for early treatment are very important based on images taken from different technologies such as CT scanner, MRI [18], X-ray [19].

In the problem of identifying and classifying breast cancer, largest sets of mammograms with information related to breast lesions are provided by experienced doctors for building a machine learning system. In particular, the mammography data reports such as Breast Imaging_Reporting & Data System_BIRADS with 7 points from BIRADS_0 to BIRADS_6 are used with the machine learning systems for evaluation. With this evaluation, testing can be utilized to help doctors to reach a more exact conclusion related to breast cancer [20].

Currently, there are many researches using Convolutional Neural Networks (CNNs) for classifying breast lesions, in which breast cancer detection is very important [21]. In addition, the CNNs are applied for classifying benign and malignant tumors on digital images with the accuracy of up to 91%. Another CNN, called Convolutional Neural Network Improvement for Breast Cancer Classification (CNNI-BCC), was structured in 2019 and applied for classifying 3 lesion cases such as benign, malignant, and normal with its accuracy of 89.5%, 90.5% and 90.7%, respectively [22].

X-ray images of diseases related to breast lesions are often of inconsistent quality and also contain many unnecessary artifacts. Therefore, in most researches, image preprocessing algorithms are applied to uniformly improve image quality and enhance the effectiveness of classification. Thus, image preprocessing before applying for training advanced deep learning networks such as VGG16 [23], EfficientNet [24], or others can increase the effectiveness of classification performance. In this research, the extract largest ROI image normalization and enhancement will be employed before using training the VGG16 deep learning network for classification of breast lesions.

This article is to evaluate the effectiveness of image preprocessing with largest ROI and enhancement before training a deep learning network of VGG16. Moreover, four types of breast x-ray images (Benign Mass, Benign Calcification, Malignant Calcification, Malignant Mass) were obtained from CBIS-DDSM [25] and MIAS [26], in which data augmentation was suggested in this research so that the classification effectiveness is increased using VGG16. In particular, image datasets have been extracted ROI, enhanced and other processing methods are applied, including rotating, flipping images and their brightness changes. It is obvious that the classification performance is increased and the obtained results are also compared to previous works for evaluation of the proposed algorithm.

2. Materials and methods

In this article, the method of a process for classification of breast lesion is proposed as described in Fig. 1. In particular, datasets are obtained from CBIS-DDSM [25] and MIAS [26] and augmentation image datasets using rotating, flipping and brightness change, in which four types of lesion images (Benign Mass, Benign Calcification, Malignant Mass, Malignant Calcification) are

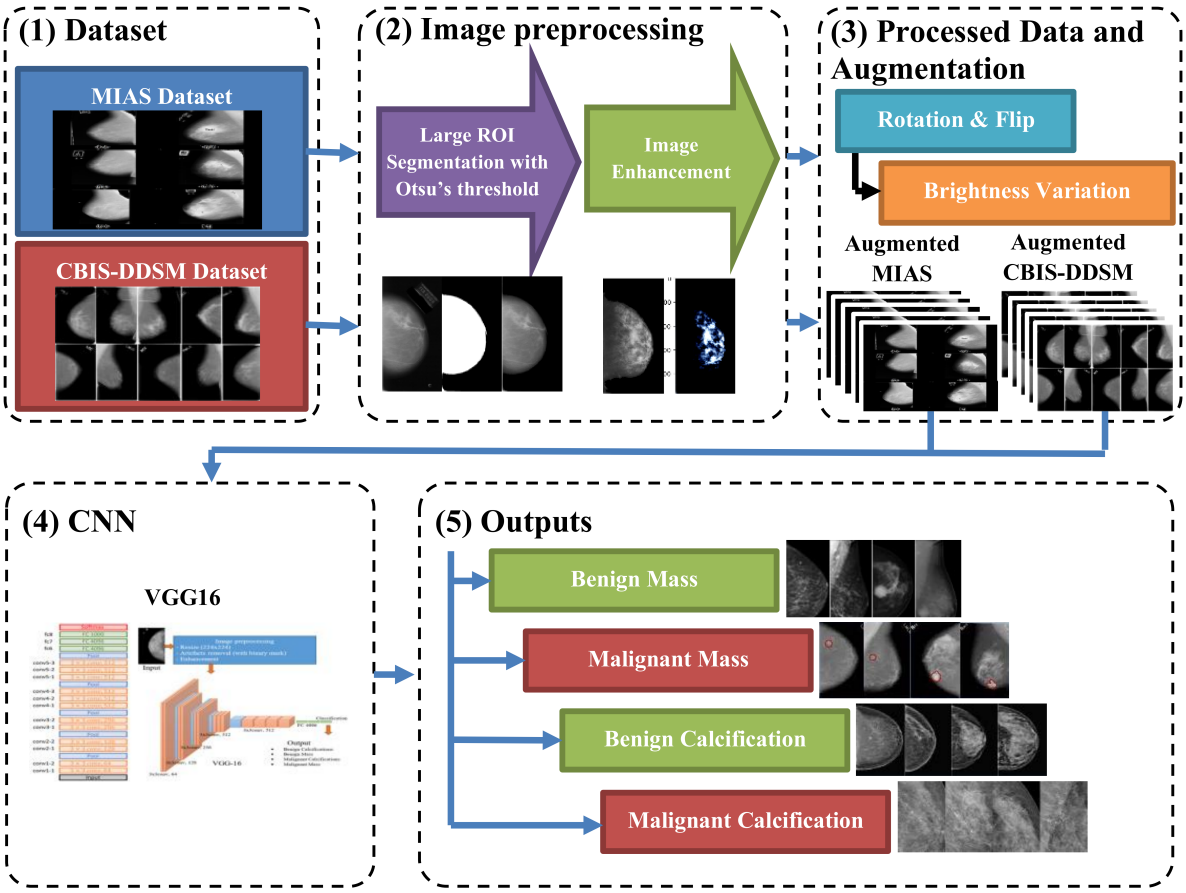


Fig. 1: Block diagram of the classification and process: (1) two datasets with four breast lesion image sets; (2) Images segmented to find the largest ROI and enhanced; (3) Images augmented from the largest ROI images; (4) and (5) the VGG16 with four outputs for breast lesion classification.

used. All image datasets are preprocessed before Otsu segmentation to choose the largest ROI containing features in each lesion image. Therefore, the largest ROI datasets are fed into a VGG16 deep learning network. This network is configured to refine the last two layers and also added appropriate output layers. Finally, the Confusion matrix is applied to evaluate the effectiveness of the proposed classification.

In addition, this article performs three cases of input image sets to the input of the VGG16 network for evaluating accuracy between them. In particular, the image sets are used for three cases such as the image sets without the Otsu segmentation, ones with just the Otsu segmentation, and the sets with the largest ROI using the Otsu segmentation.

2.1. Materials

Mammograms are used in this research from two image sources: CBIS-DDSM and MIAS. Moreover, CBIS-DDSM is an updated version of the DDSM providing easily accessible data and improved ROI segmentation [25]. The original DDSM is a database of 2,620 scanned film mammography studies, while the CBIS-DDSM has 1,644 images of four types of breast lesions (Benign Mass; Malignant Mass; Benign Calcification; Malignant Calcification) and the image number in each type and their information are shown in Tab. 1. In similarity, the MIAS Mammography dataset includes 1,128 images of 362 women and only 322 images of four types of breast lesions with their information and the image number are described in Tab. 2.

Tab. 1: Information of the CBIS-DDSM dataset.

Properties	Value
Total Number of Images	1644
Image Dimension	224 × 224
Color Codec	RGB
Benign Mass	753
Benign Calcification	414
Malignant Calcification	339
Malignant Mass	419

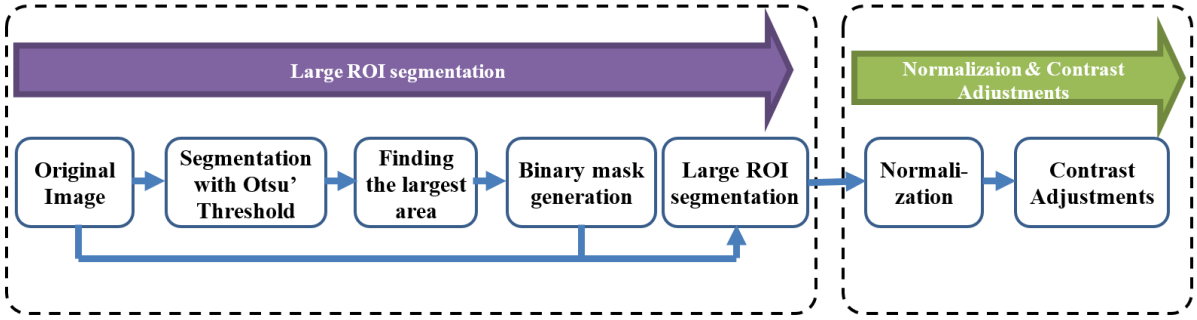


Fig. 2: Description of process for creating the largest ROI with enhancement.

Tab. 2: Information of MIAS dataset.

Properties	Value
Total Number of Images	322
Image Dimension	224 × 224
Color Codec	RGB
Benign Mass	37
Benign Calcification	12
Malignant Calcification	13
Malignant Mass	19

2.2. Determination of the largest ROI

In this research, all images are processed to determine the largest ROI focused on the breast lesion area before used to train the VGG16 network for classification of breast lesion images. With processing the images to just produce the largest ROI, the accuracy classification is high and also the classification system reduces the model's computation time. In particular, the Otsu algorithm is applied to an original image to determine a threshold for image segmentation to obtain objects in the image. Moreover, the largest ROI with the breast lesion part is determined based on the largest pixel number and the smaller ROI parts are eliminated. This image with the largest ROI is a binary image which is multiplied with the original image to produce the original image with the largest ROI only. Finally, the largest ROI image is normalized and adjusted to produce the wanted ROI image as shown in Fig. 2.

An original mammogram may contain unwanted components which do not contribute for accurate image classification, but also reduce classification accuracy. Furthermore, these components can increase training and recognition times for classifiers. In particular, lines of text or notes on the image or bright areas at the edges of the image can affect the training and classification process. To increase classification accuracy, mammograms need to be processed to remove these areas. Furthermore, the background component of the image will be removed to determine the largest ROI before entering the network for training and classification.

Image segmentation is the process of dividing an image into regions with similar features. In addition, edges and textures are important features which need to obtain for segmentation. In this article, gray mammograms are segmented based on thresholding using the Otsu method. Suppose the threshold $q = 1, 2 \ldots, (L-1)$ is chosen to divide the number of gray levels L into two sets C_0 and C_1 and the Otsu threshold method allows to find q based on the largest variance between the two sets C_0 and C_1 and its equation is expressed as follows:

$$\sigma_B^2 = \omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2 \tag{1}$$

in which μ_T is the average gray level of one image.

In order to determine this variance, the weight ω_0 and the average gray level μ_0 of the set C_0, the weight ω_1 and the average gray level μ_1 of the set C_1 are calculated as follows:

$$\omega_0 = \sum_{q=0}^{k-1} p_q(r_q) \tag{2}$$

$$\omega_1 = \sum_{q=k}^{L-1} p_q(r_q) \tag{3}$$

$$\mu_0 = \sum_{q=0}^{k-1} q p_q(r_q) / \omega_0 \tag{4}$$

$$\mu_1 = \sum_{q=k}^{L-1} q p_q(r_q) / \omega_1 \tag{5}$$

Therefore, the probability density p_q of the gray level is calculated as follows:

$$p_q(r_q) = \frac{n_q}{M \times N} \tag{6}$$

in which, n_q is the pixel number of the q^{th} gray level and $M \times N$ is the pixel number of one gray image.

The Otsu segmentation method will give many different thresholds, but to choose a suitable threshold, one is based on the variance value of the two sets as in Eq. (1). In particular, the largest variance is the value considered to obtain the best threshold for segmentation. However, a threshold close to the recommended threshold can be chosen for the best segmentation and it is suitable for the typical data.

To effectively remove unwanted components and obtain only the largest ROI, an algorithm to find the largest region based on pixels is applied. In particular, the gray segmented image will be converted to the binary image for determining the largest ROI based on the white pixels, while eliminating smaller regions. Furthermore, to obtain the original image with only the largest ROI, the binary image with the largest ROI will be multiplied with the original image.

For determining the largest ROI, an Otsu threshold is applied to produce objects with the same pixel gray levels and then a largest object, called the largest ROI, will be found. In particular, the algorithm for finding the largest ROI will be determined as follows:

$$\chi_i = \sum_{x=0}^{h_i} \sum_{y=0}^{w_i} O_i(x, y) \tag{7}$$

in which, χ_i is the sum of the white pixels, O_i is the i^{th} ROI with the size of $h_i \times w_i$ $(i = 1, 2, \ldots, n)$.

After completing the largest ROI segmentation, the images are normalized and enhanced using histogram equalization. The enhancement will make all synchronous images with features and brigthness and the classification will be more accurate. The images after enhancement will be fed into the VGG16 network for training and classification.

To normalize images, the method with the values of min and max gray levels is applied in this study. Therefore, the new gray level value is normalized using the following equation:

$$x_{new} = x_{old} \frac{x_{old} - x_{\min}}{x_{\max} - x_{\min}} \tag{8}$$

in which, x_{new} is the new gray level of pixel. x_{old} is the old gray level of pixel. $x_{\max}$ describes the maximum gray level of image. And $x_{\min}$ describes the minimum gray level of image.

To enhance the image, the histogram method is utilized. In particular, this method will recalculate the gray level values based on the probability density of each gray level, in which it has the relationship between the probability density of the next gray level and that of the previous gray level through sum calculation. Therefore, each gray level is calculated for equlization based on histogram in the image described by the following equation:

$$s_q = (L-1) \sum_{q=0}^{L=1} p_q(r_q) \tag{9}$$

in which, L is the number of the gray level with $q = 0, 1, 2, \ldots, L-1$.

Assume that r_q is the gray level value in the original and s_q is the gray level in the equalized image and the probability density p_q of the gray levels of the original image as in Eq. (6). Furthermore, nq is the number of pixels at the gray level $q = 0, 1, 2, \ldots, L-1$ and $M \times N$ is the number of pixels horizontally and vertically in an image.

2.3. VGG16 network model

VGG16 is a model of a Deep Convolutional Neural Network (DCNN) [19]. This model achieved 92.7% accuracy in the top-5 of testing in the ImageNet dataset [23]. The increased depth of the VGG model can support kernels for extracting more complex features. In this article, VGG16 was used to classify four types of breast lesions (Benign Mass; Malignant Mass; Benign Calcification; Malignant Calcification) of breast X-ray image sets with lesions. In addition, the VGG16 network is fine-tuned so that it can achieve higher accuracy compared to the full network as depicted in Fig. 3. Furthermore, the computer configuration was used to train the VGG16 network includes an Intel Core i9-9980HK processor, NVidia Tesla P4 GPU, 32GB DDR4 Ram Memory, and 1TB SSD for storage. With this configuration, the network performs relatively well and gives more accurate classification results.

In this article, a VGG16 with the total 39 layers and each layer has five blocks. Moreover, each block contains two convolutional layers followed by a Maxpooling layer. Therefore, each convolution layer is multiplied to a two-dimensional array with weights of input data for performing a linear operation. The input layer of the architecture requires the size of the image with (224×224×3) and the input size for the first convolutional layer is similar. In particular, the first block has two convolutional layers with 64 channels with the kernel size (3×3) and the same padding, followed by a (2×2) Maxpooling layer. In similarity, the second block contains two convolutional layers of 128 channels with th 3×3 kernel and they are followed by a Maxpooling layer with the (2×2) size. The last three blocks contain three convolutional layers with the Maxpooling layer. In addition, the channel sizes of the three convolutional layers in blocks 3, 4, and 5 are 256, 512, and 512, respectively and all of them are the (3×3) kernel size. Thus, the original input image is scaled down to half its size within each Maxpooling layer. After stacking the convolutional and Maxpooling layers, the feature output from the final Maxpooling layer has a size of (7×7×512). A flattening layer is added to create a (1×25,088) features vector. Moreover, a dense layer is added to produce four channels for the layers. In this deep learning network, there is a Softmax activation function at the end for normalizing the classification vector obtained from FC. Finally, the VGG16 model is pre-trained on the ImageNet database.

The convolutional layers in the VGG16 model use the ReLU loss function, which is equal to 0 for $x < 0$

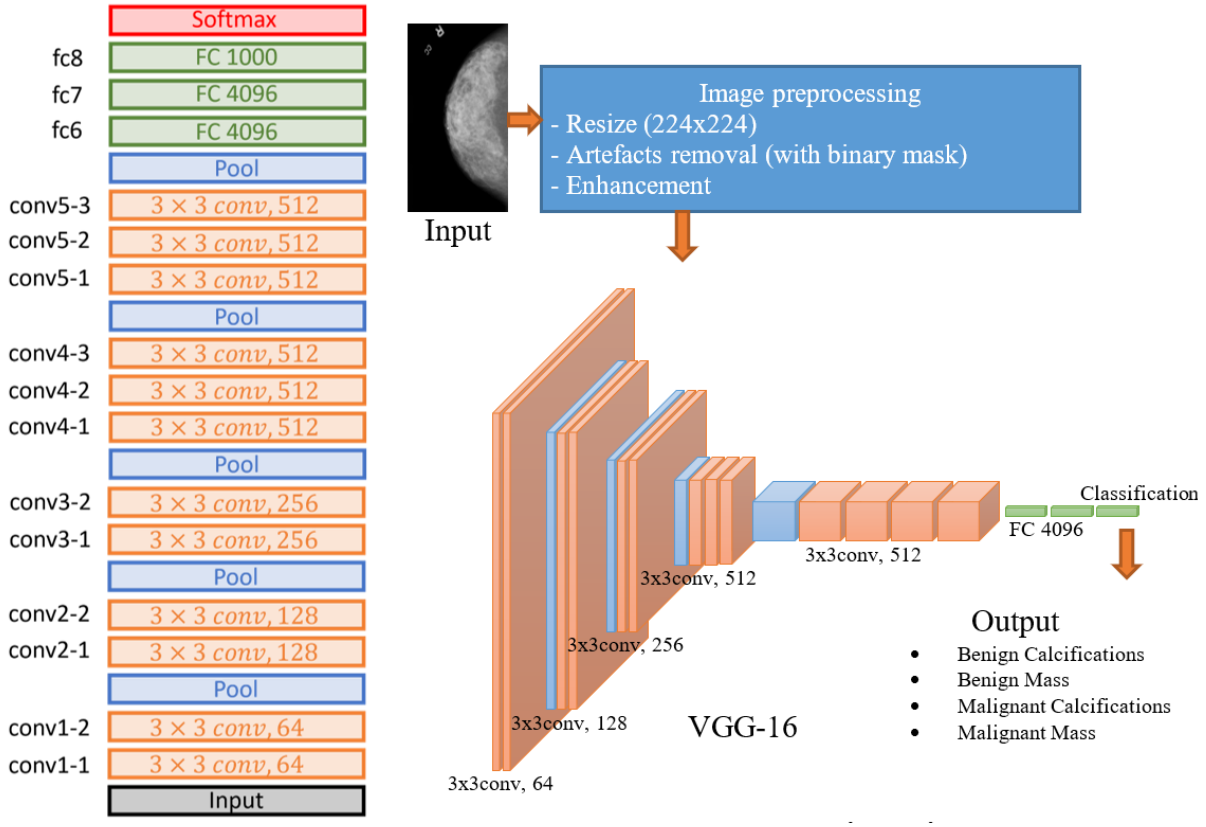


Fig. 3: Architecture of the fine-tuned VGG16 network used in this article.

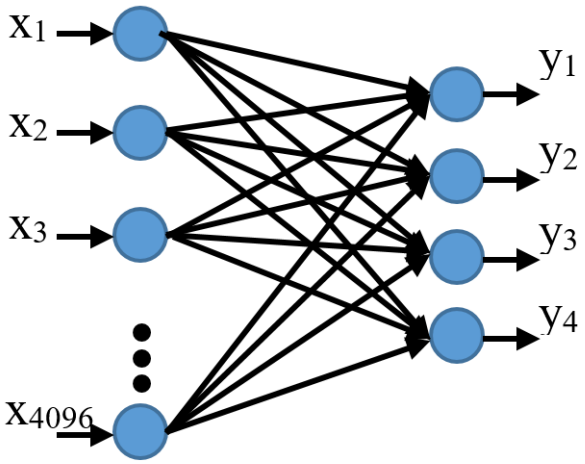


Fig. 4: Description of the last FC layer.

and equal to x for all $x \geq 0$ and it is described as follows:

$$ReLU\,(x) = \max\{x, 0\} = \begin{cases} x, & if \quad x \geq 0 \\ 0, & if \quad x < 0 \end{cases} \tag{10}$$

In this article, the VGG16 model has two phases: training and fine-tuning. During training, the last two layers of 16 layers in the VGG16 model are fine-tuned, and weights in the remaining layers are kept from the prototype trained on ImageNet. In the output layer of the model, the dense and FC layers of the network model are updated by the new features of the images during the fine-tuning process. Moreover, the final FC layer will be modified for classifying four types of breast lesion images. In particular, to achieve the purpose of classifying 4 breast lesion image layers including Benign Mass, Benign Calcification, Malignant Calcification, and Malignant Mass, the last FC layer is adjusted to have 4 outputs as shown in Fig. 4.

Therefore, the pre-trained weights of layers are transferred by using transfer learning and then the model is retrained using the image datasets. In this model, the 'Softmax' activation function is used to calculate the probability for each class and to create the prediction results. Therefore, Softmax is expressed by the following equation:

$$Z_j = \frac{e^{x_j}}{\sum_{i=1}^{n} e^{x_i}}; \quad j = 1, 2, \ldots, n \tag{11}$$

To train the model in this research, the maximum number of epochs is set to 350 and the batch size to 16. The Adam optimization algorithm is applied with a learning rate of 0.001. In addition, the Adam algorithm is an enhanced version of Stochastic Gradient Descent, which is a combination of the Momentum and RMSP algorithms and is described by the following mathe-

matical equation:

$$W_{T+1} = w_t - \propto m_t \tag{12}$$

in which,

$$m_t = \beta m_{t-1} + (1-\beta)\left[\frac{\delta L}{\delta w_t}\right] \tag{13}$$

and, W_{T+1} is weights at time t+1. w_t is weights at time t. $\propto$ is learning rate. m_t is aggregate of gradients at time t. δL is derivative of Loss Function. δw_t is derivative of weights at time t. And β is moving average parameter.

3. Results and discussions

This article proposes an algorithm for the classification of breast lesions, in which image sets are processed to extract the largest ROI before inserting them into VGG16. In this algorithm, the results of image processing and classification using VGG16 are demonstrated to illustrate the effectiveness of this algorithm. Finally, the comparison and evaluation of the results are performed to illustrate that the proposed method is effective.

3.1. Determination of largest ROI

With a breast lesion image, it is very necessary to accurately determine whether it is breast cancer because this can help doctors have more information to decide breast cancer. In this study, image datasets with breast lesions were segmented to extract the largest ROI parts with many features for better lesion classification using deep learning networks. In particular, to remove unwanted components in the images for creating the largest ROI, the image segmentation method using Otsu thresholding was utilized as shown in Fig. 5.

Segmentation with different threshold values can produce segmented images with different ROI parts. Therefore, simulation results in this research showed that adjusting image brightness before segmentation significantly improved the regions segmented with the Otsu algorithm. With a chosen Otsu threshold, extracting the largest ROI is obtained. In particular, the breast lesion image only retains the largest ROI containing many possible features and most of the unwanted components in the original image could be eliminated as shown in Fig. 5.

In Fig. 5, the images of (a4), (b4), (c4), (d4) just have the largest ROI and smaller ROI parts are eliminated compared to the images (a1), (b1), (c1), (d1). Moreover, the images of (a4), (b4), (c4), and (d4) after eliminating can be seen more clearly lesions, called features, around ROI. The processed images make it easy to see the difference between the four types of breast lesions (BM, MM, BC, MC).

Fig. 6 shows the ROI image normalized and adjusted for contrast using the Histogram equalization algorithm. In particular, the images after correction were clearer with features focusing on the largest ROI parts and this can help the cancer classification system achieve better performance. In addition, the enhanced images of (a3), (b3), (c3), and (d3) produced the largest ROI parts nearly synchronous, clearer compared to the images of (a1), (b1), (c1), (d1). This is important due to the ability to classify with higher accuracy using VGG16 compared to the images without enhancement.

3.2. Determination of largest ROI

In this article, two main datasets of CBIS-DDSM and MIAS are used for the classification of four breast lesions using the deep learning network. However, they are not enough for training and classifying to produce the desired accuracy. Thus, an augmentation of image sets is necessary for using the deep learning network. The rotation and flip images in these two datasets and also the brightness adjustment applied were performed as shown in Tab. 3 and Tab. 4.

All image sets after the largest ROI extraction and enhancement are used for training the network. In this article, the VGG16 is applied for the classification of four types of breast lesions. To evaluate the performance of training, the datasets were divided into the training set (80%) and the testing set (20%), respectively. In particular, after data augmentation, 15,118 X-ray images of breast lesions were extracted with the largest ROI divided into 12,093 for training (80%) and 3,025 images for testing (20%) as depicted in Fig. 7. In particular, the average class for training is around 3,000 images and similarly it is 750 images for testing in this research.

3.3. Results for classification of breast lesion images

Fig. 8 shows the training curves for the classification model, including the loss curve and the accuracy curve. It is obvious that the training curve smoothly converges from the first stage to the last stage without collisions. Furthermore, the distance between the validation accuracy curve (red) and the training one (blue) shows that they nearly have no overfitting during the training process. In similarity to the training curve, the loss curve shown in Fig. 8 converges steadily. With these

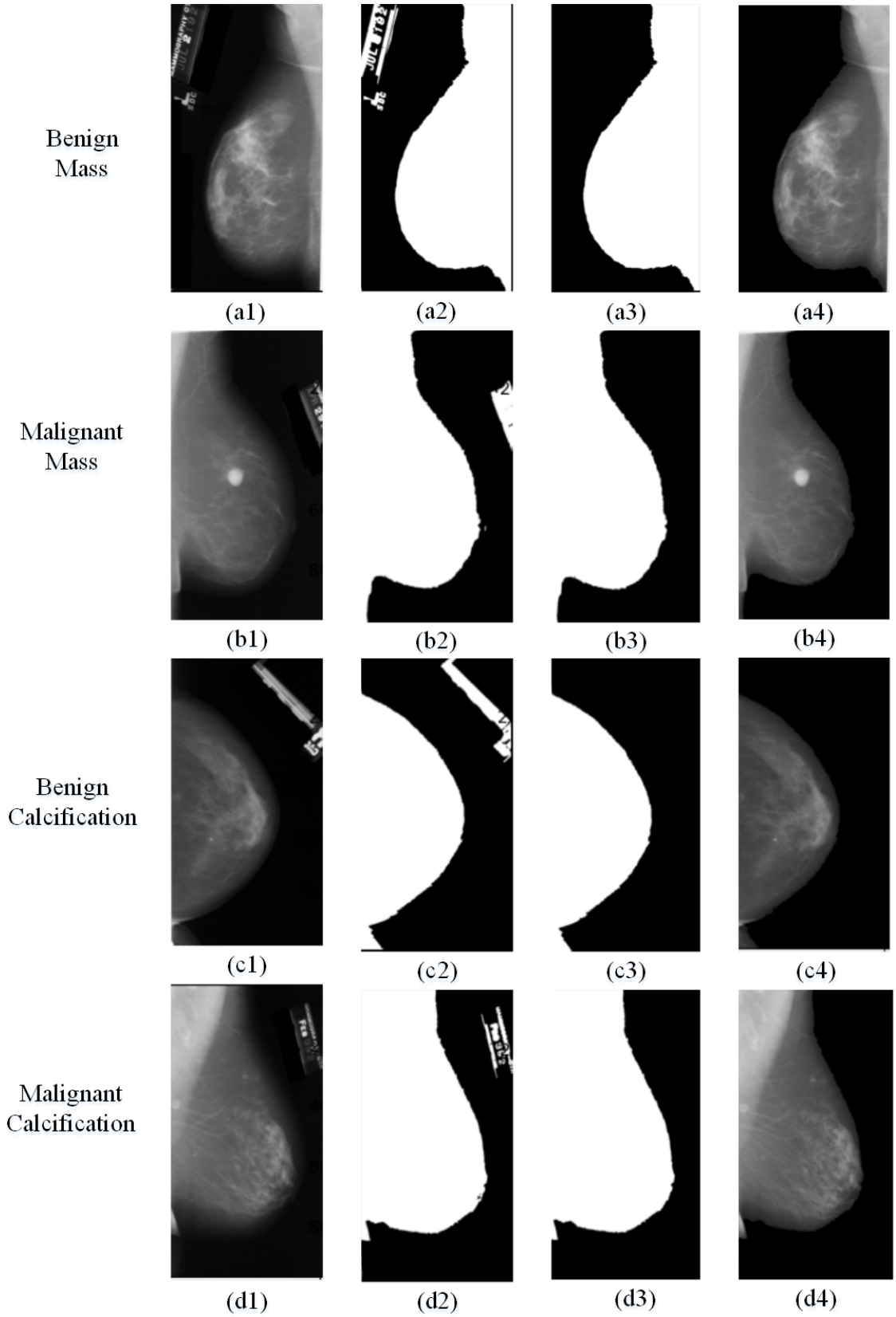


Fig. 5: Description of extracting the largest contour-based ROI and eliminating unwanted components: (a1-a4) Original images of four breast lesions; (b1-b4) Segmented images with the Otsu threshold method (c1-c4) Binary images with the largest ROI eliminated unwanted components; (d1-d4) Original image with the largest ROI without unwanted components.

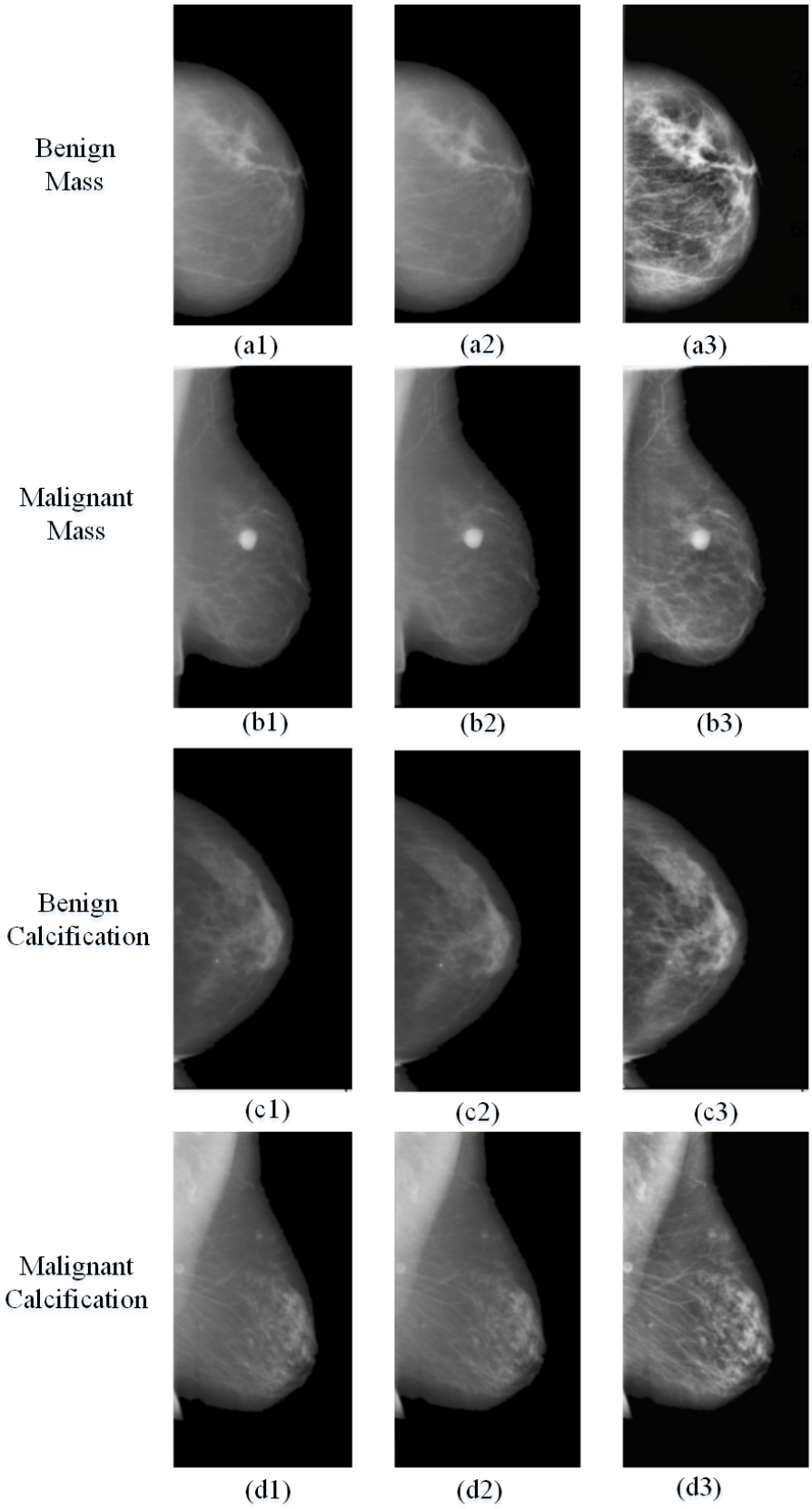


Fig. 6: Description of results for normalizing and the largest ROI image enhancement using histogram equalization: (a1-a3) Original images; (b1-b3) Normalized images; (c) Image with the histogram equalization.

Tab. 3: Description of four types of the augmented image sets of CBIS-DDSM.

	Original images	Augmented images	The augmented sets divided for training and testing	
			Training	Testing
Image No.	1644	12542	10033	2509
Benign Mass	753	3165	2532	633
Benign Calcification	414	2897	2318	579
Malignant Calcification	339	2985	2387	598
Malignant Mass	419	3495	2796	699

Tab. 4: Description of four types of the augmented image sets of CBIS-DDSM.

	Original images	Augmented images	The augmented sets divided for training and testing	
			Training	Testing
Image No.	322	2576	2060	516
Benign Mass	146	1168	934	234
Benign Calcification	48	384	307	77
Malignant Calcification	52	416	333	83
Malignant Mass	419	3495	2796	699

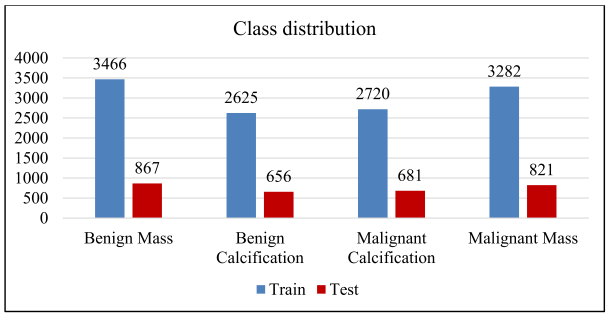


Fig. 7: Distribution of four types of datasets for training (blue) and validation (red).

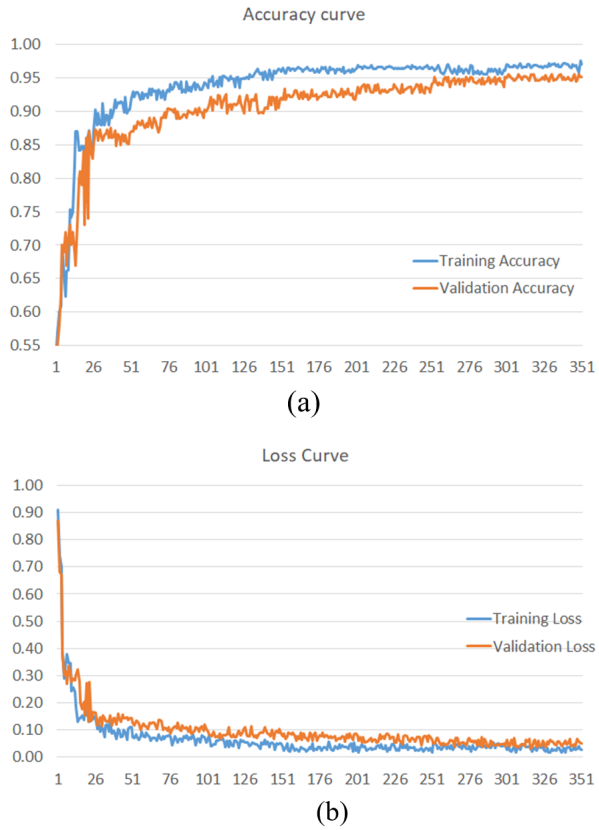


Fig. 8: Description of training process with over 350 epochs for Adam with learning rate 0.0001: (a) accuracy curve for training and validation; (b) loss curve for training and validation.

results of the training and loss curves without overfitting or underfitting, we can see that the proposed model for the classification of four breast lesions is effective. In addition, the accuracy rate over the epochs curve shows that the accuracy increases when the number of iterations increases, in which the blue line is for the training set, and the red line is for the test set.

Fig. 9(a) and Fig. 9(b) show the results of using the confusion matrix for the VGG16 classification model with the highest accuracy. In particular, the row values represent the actual labels of the image sets, BC, BM, MC, and MM, the column values describe the predicted labels of four types of these image sets, and the diagonal lines express the True Positive (TP) values. In addition, the confusion matrix shows that no bias exists for any layer and also, they predict all similar layers. Specifically, with the CBIS-DDSM dataset, the model can classify MM and BM lesions with the highest accuracies of 95.71% and 95.58% respectively. Meanwhile, with the MIAS dataset, BM and BC lesions yield the highest classification accuracies of 93.16% and 92.21%

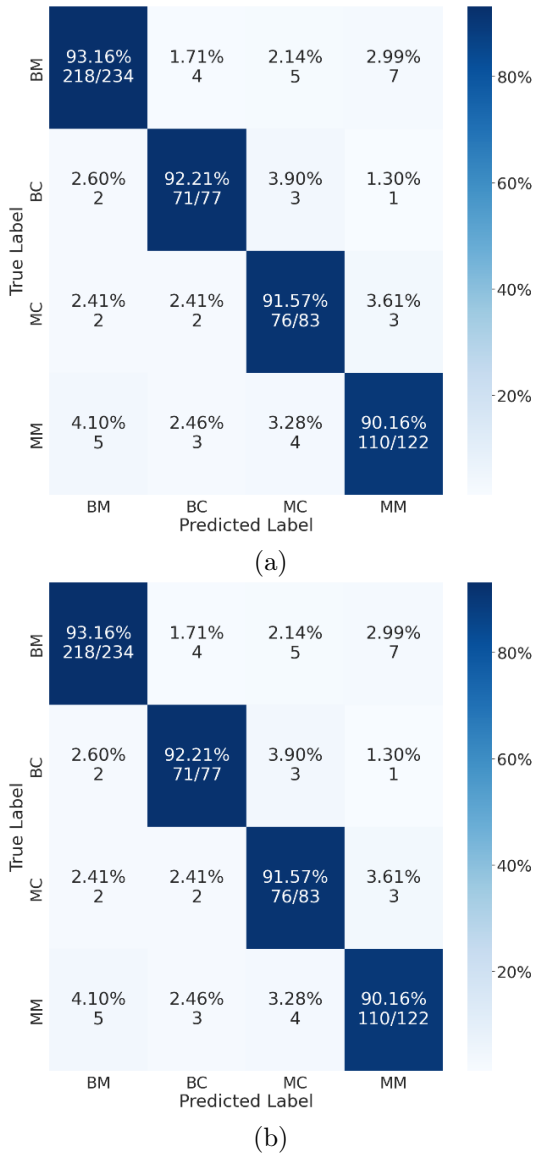


Fig. 9: Description of training process with over 350 epochs for Adam with learning rate 0.0001: (a) accuracy curve for training and validation; (b) loss curve for training and validation.

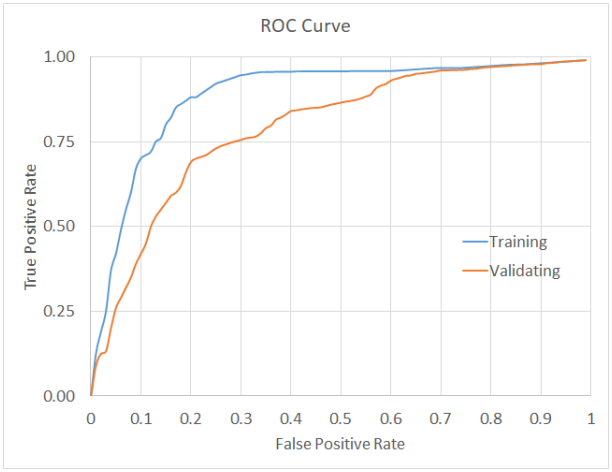


Fig. 10: Accuracy rate curves of training and testing using the proposed model.

respectively. The average accuracy over the CBIS-DDSM dataset is 95.23%, which is higher than the accuracy achieved when classifying on the MIAS dataset at 91.78%. Notably from Fig. 9, the misclassification rates from BM to MM are high for both test datasets at 1.74% and 2.99% respectively. Additionally, MM is also misclassified as BM with rates of 2.00% and 4.10% on the CBIS-DDSM and MIAS datasets. These classification results demonstrate the effectiveness and potential of the proposed model.

In Fig. 10, the accuracy curves for testing and validating show that the model achieves high accuracy with a low false positive rate for the data sets. In particular, the yellow curve is the result of the training set and the blue curve is the result of the testing set. Thus, the receiving operating curve (ROC) and the area under the curve (AUC) taken from the ROC curve show that the proposed model along with the proposed method is effective for distinguishing different classes. In particular, an AUC value close to 1 indicates that this model can detect most classes, where the ROC curve almost touches the peak of the y-axis, which means the false positive rate is close to 0. The true positive rate is close to 1, which proves that the model's effectiveness is very high.

Extracting the largest ROI and enhancing images has improved the classification performance of the proposed model, particularly, the proposed model in this article achieved 95% accuracy. For the high performance, the image sets were extracted with the largest ROIs to capture many features of breast lesions and then normalized and enhanced. For the effective evaluation of the proposed model, this article shows three cases of data sets as shown in Tab. 5. In particular, the image set is processed following the proposed with the largest ROI producing the highest performance of 95%, while the image set without processing is 82.1% of accuracy smaller than that of just Otsu segmentation. This result proves that the proposed model can classify four types of breast lesions to make it easier to identify breast cancer from mammograms with the highest accuracy. One outstanding point is that applying the large ROI region segmentation based on the Otsu algorithm allows for quick segmentation without using pre-segmented data for training. Furthermore, compared to other deep learning networks, such as U-Net, for segmented data, this method is simpler, and less time-consuming, but still increases the average accuracy of the system.

To evaluate the reliability of the results, with a limited number of data sets, we use the cross-validation method with 5-fold on the same data set. Fig. 11 illus-

Tab. 5: Information of the CBIS-DDSM dataset.

Methods	Average Accuracy
No image processing	82.1%
Otsu segmentation	89.2%
Otsu segmentation with the largest ROI	95%

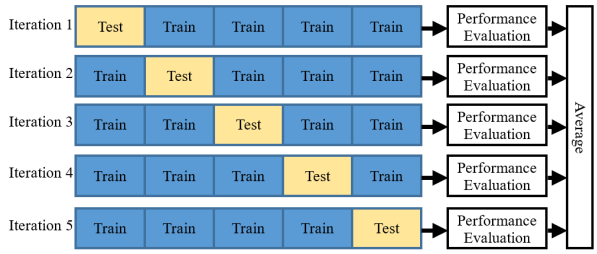


Fig. 11: 5-fold cross validation.

trates how to divide the data set for use in this evaluation process. In particular, the data set is divided into 5 segments, each iteration uses 1 of the 5 segments in turn as the validation set, while the remaining 4 segments are used for training. The results of 5 iterations are used to compare with each other for evaluating the stability of the model.

With the results of the 5 iterations as shown in Tab. 6, there is no big difference between the iterations and these results are reliable. Moreover, from these results, Confidence Intervals (CIs) are calculated and shown in Fig. 12. In particular, the CI of Accuracy is quite narrow, showing the stability of accuracy. The confidence intervals of Sensitivity, Precision, and F1 score are larger but still do not have too large deviations between iterations. These results show the stability of the model.

Tab. 7 shows the highest accuracy achieved by the proposed model compared to previous works with similar methods, datasets, or models. More than half of

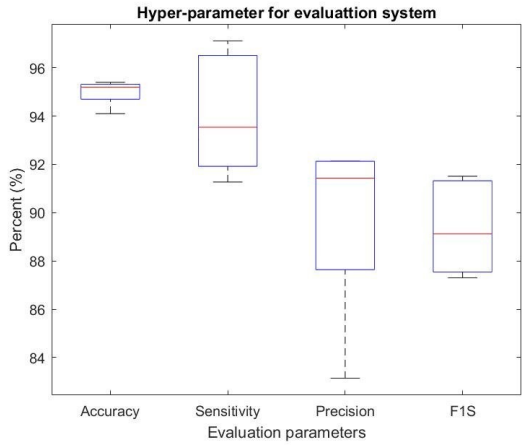


Fig. 12: Mean value and confidence interval of Accuracy, Sensitivity, Precision, and F1s.

the previous studies in Tab. 7 used the VGG16 and similar datasets compared to the study in this article. In [27], the transfer learning method with the MVGG model was proposed for image classification based on the VGG16 model and the result was 94.3% with 15 epochs using the DDSM database.

In a study by Shallu and Mehra [28] using the pre-trained and fine-tuned VGG16, its performance achieved the highest accuracy of 92.60%. With VGG16 and VGG19s using the same dataset, Hameed et al. [29] had the result with the obtained accuracy of 94.71%. Similarly, in the research of Li et al. [30], his article applied several pre-trained models and the highest accuracy was 92.78% with VGGNet. Our fine-tuned VGG16 network with the FC layer achieved 95% accuracy in breast lesion classification using the CBIS-DDSM dataset after Adam optimization. In general, with the above statistics, the models using VGG16 achieved the classification with a very high accuracy.

Two image sets (CBIS-DDSM and MIAS) are very popular for research related to breast lesions. Although different algorithms of classification, the accuracy is very high, around 92%. In our research, these two image sets with four types of breast lesions were used, in which there were the augmented image sets to create the large dataset. In addition, these image sets were processed to extract the largest ROI, which was used for the fine-tuned VGG16 and produced the results with a high accuracy of 94% to 95%. This means that the obtained results demonstrate to illustrate the effectiveness of the proposed algorithm.

4. Conclusion

This article has proposed a classification model using four breast lesion image sets, including Benign Calcifications (BC), Benign Mass (BM), Malignant Calcifications (MC), and Malignant Mass (MM) from two image sets of CBIS-DDSM and MIAS. All images were extracted with the largest ROI and then normalized and enhanced before applying to train the VGG16 network. Extracting the largest ROI gave a higher classification efficiency of 95% compared to images without extracting the largest ROI. Furthermore, to increase accuracy, these image sets were processed to augment the number of images by rotating and flipping images. In addition, in VGG16, the last two layers are fine-tuned during training and classification to contribute to the classification performance. Therefore, the results demonstrate to illustrate that removing unwanted components and extracting only the largest ROI significantly contributes to improving the proposed model performance. The proposed model with VGG16 shows high performance for classifying four types of breast lesions of CBIS-DDSM and MIAS sets. With the clas-

Tab. 6: Accuracy, Sensitivity, Precision, and F1s over 5 iterations.

Iteration	Accuracy	Sensitivity	Precision	F1s
1	95.30%	0.9631	0.9214	0.8763
2	95.40%	0.9712	0.9213	0.8912
3	94.10%	0.9354	0.8914	0.9126
4	94.90%	0.9127	0.9143	0.8730
5	95.20%	0.9214	0.8314	0.9151
Average	**94.98%**	**0.9408**	**0.8960**	**0.8936**

Tab. 7: Comparison between this research with previous work.

Paper	Model	Dataset	Epoch	Accuracy
Khampari et al. (2021) [27]	Hybrid MVGG16 ImageNet	Digital Database with DDSM having 2620 images.	15	94.3%
Shallu and Mehra, (2018) [28]	VGG16 + LR VGG19 + LR ResNet50 + LR	2042 full-field digital images provided by the 1st Hospital of Shanxi Medical University and 30,630 augmented images	-	92.60% 90.40% 79.40%
Hameed et al. (2020) [29]	Full trained VGG16 + VGG19 Fine-tuned VGG16 + VGG19	544 images provided by Colsanitas Colombia University	200	93.53% 95.29%
Li et al. (2019) [30]	AlexNet VGGNet GoogLeNet DenseNet DenseNet-II	2042 full-field from the First Hospital of Shanxi Medical University and augmentation of 30,630 images	-	92.70% 92.78% 93.54% 93.87% 94.55%
Al-antari et al. (2018) [30]	Fully Integrated CAD	410 Full-field Digital images and 896 augmented images	100	95.64%
Nasir Khan et al. (2019) [31]	MVFF CADx	3568 CBIS-DDSM images and 322 MIAS mammograms	100	93.73%
Zahra Jafari and Ebrahim Karami (2023) [32]	Concat. Model with the NN classifier	54,713 images of RSNA dataset	-	96%
This study	Largest ROI and fine-tuned VGG16	12,093 images of CBIS-DDSM including augmentation and 3,025 augmented images of MIAS	350	95%

sification performance up to 95%, the proposed model may be developed for classification with many more different breast lesion image sets to support in breast cancer diagnosis. In addition, with the results of this research, the proposed method can be developed to be able to apply to other deep learning networks and different datasets.

Acknowledgment

We would like to thank Ho Chi Minh City University of Technology and Education (HCMUTE), Vietnam.

Author Contributions

Thanh-Tam NGUYEN developed the theoretical formalism, performed the analytic calculations, and performed the numerical simulations to produce results.

Thanh-Hai NGUYEN supervised and edited the whole article. Ba-Viet Ngo contributed to the draft version of the manuscript and evaluated the results. Thanh-Nghia NGUYEN contributed to the acquisition of data and augmented and analyzed them.

References

[1] ABAAN, O. D., W. E. CRISS. Gene Therapy in Human Breast Cancer. *Turkish Journal of Medical Sciences*. 2002, vol. 32, pp. 283–291. https://journals.tubitak.gov.tr/medical/vol32/iss4/1/.

[2] FACKENTHAL, J. D., O. I. OLOPADE. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. *Nat Rev Cancer*. 2007, vol. 7, no. 12, pp. 937–48. DOI: 10.1038/nrc2054.

[3] FACKENTHAL, J. D., O. I. OLOPADE. Telemammography: A Novel Approach for Early Detection of Breast Cancer Through Wavelets Based Image Processing and Machine Learning Techniques. *Advances in Soft Computing and Machine Learning in Image Processing*. 2018, pp. 149—183. DOI: 10.1007/978-3-319-63754-9_8.

[4] ARNOLD, M., et al. Current and future burden of breast cancer: Global statistics for 2020 and 2040. *The Breast*. 2022, vol. 66, pp. 15–23. DOI: 10.1016/j.breast.2022.08.010.

[5] ARNOLD, M., et al. Comparison of wire-guided localization and radio-guided occult lesionlocalization in preoperative localization of nonpalpable breast lesions. *Turk J Med Sci*. 2016, vol. 46, no. 6, pp. 1829–1837. DOI: 10.3906/sag-1507-162.

[6] YAN, J., Z. LIU, S. DU, J. LI, L. MA, L. LI. Diagnosis and Treatment of Breast Cancer in the Precision Medicine Era. *Methods Mol Biol*. 2020, vol. 2204, pp. 53–61. DOI: 10.1007/978-1-0716-0904-0_5.

[7] ZHANG, J., J. HU. Image Segmentation Based on 2D Otsu Method with Histogram Analysis. *2008 International Conference on Computer Science and Software Engineering, Wuhan, China*. 2008, pp. 105–108. DOI: 10.1109/CSSE.2008.206.

[8] MUSTRA, M., M. GRGIC, R. M. RANGAYYAN. Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms. *Medical & Biological Engineering & Computing*. 2016, vol. 54, pp. 1003–1024. DOI: 10.1007/s11517-015-1411-7.

[9] SHI, P., J. ZHONG, A. RAMPUN, H. WANG. A hierarchical pipeline for breast boundary segmentation and calcification detection in mammograms. *Computers in Biology and Medicine*. 2018, vol. 96, pp. 178–188. DOI: 10.1016/j.compbiomed.2018.03.011.

[10] KASHYAP, K. L., M. K. BAJPAI, P. KHANNA, G. GIAKOS. Mesh-free based variational level set evolution for breast region segmentation and abnormality detection using mammograms. *International Journal for Numerical Methods in Biomedical Engineering*. 2018, vol. 34, iss. 1. DOI: 10.1002/cnm.2907.

[11] CELIK, T., T. TJAHJADI. Automatic Image Equalization and Contrast Enhancement Using Gaussian Mixture Modeling. *IEEE Transactions on Image Processing*. 2012, vol. 21, pp. 145–156. DOI: 10.1109/TIP.2011.2162419.

[12] YAO, L., S. MUHAMMAD. A novel technique for analysing histogram equalized medical images using superpixels. *Computer Assisted Surgery*. 2019, vol. 24, pp. 53–61. DOI: 10.1080/24699322.2018.1560100.

[13] NGUYEN, T. H., N. T. NGUYEN, M. H. NGUYEN, S. LIVATINO. Wavelet-Based Kernel Construction for Heart Disease Classification. *Advances in Electrical and Electronic Engineering*. 2019, vol. 17, no. 3, pp. 306-319. DOI: 10.15598/aeee.v17i3.3270.

[14] NGUYEN, M.-H., V.-H. TRAN, T.-H. NGUYEN, T.-N. NGUYEN. A Deep Learning Framework for Inter-Patient ECG Classification. *IJCSNS International Journal of Computer Science and Network Security*. 2019, vol. 19, pp. 74–84. DOI: 10.15598/aeee.v17i3.3270.

[15] WEI, L.-S., Q. GAN, T. JI. Skin Disease Recognition Method Based on Image Color and Texture Features. *Comput Math Methods Med*. 2018, vol. 2018. DOI: 10.1155/2018/8145713.

[16] PEREIRA, S., A. PINTO, V. ALVES, C. A. SILVA. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Transactions on Medical Imaging*. 2016, vol. 35, no. 5, pp. 1240–1251. DOI: 10.1109/TMI.2016.2538465.

[17] HUSAINI, M. A. S. A., M. H. HABAEBI, S. A. HAMEED, M. R. ISLAM, T. S. GUNAWAN. A Systematic Review of Breast Cancer Detection Using Thermography and Neural Networks. *IEEE Access*. 2020, vol. 8, pp. 208922–208937. DOI: 10.1109/ACCESS.2020.3038817.

[18] KESKIN, N. K., et al. Detection of the differences in the apparent diffusion coefficient values in different histopathological types of malignant breast lesions and comparison of cellular region/stroma ratio and histopathological results. *Turkish Journal Of Medical Sciences*. 2018, vol. 48, pp. 817–825. DOI: 10.3906/sag-1801-89.

[19] YÜCEL, A., B. DEĞİRMENCİ, M. ACAR, H. ELLİDOKUZ, R. ALBAYRAK, A. HAKTANIR. Knowledge About Breast Cancer and Mammography in Breast Cancer Screening Among Women Awaiting Mammography. *Turkish Journal of Medical Sciences*. 2005, vol. 35, pp. 35-–42. https://journals.tubitak.gov.tr/medical/vol35/iss1/6/.

[20] DOMINGUES, I., P. H. ABREU, J. SANTOS. Bi-Rads Classification of Breast Cancer: A New Pre-Processing Pipeline for Deep Models Training. *2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece*. 2018, pp. 1378–1382. DOI: 10.1109/ICIP.2018.8451510.

[21] AMALA, R. K., H. -P. CHAN, L. HADJIISKI, M. A. HELVIE, C. D. RICHTER, K. H. CHA. Breast Cancer Diagnosis in Digital Breast Tomosynthesis: Effects of Training Sample Size on Multi-Stage Transfer Learning Using Deep Neural Nets. *IEEE Transactions on Medical Imaging*. 2019, vol. 38, no. 3, pp. 686-696. DOI: 10.1109/TMI.2018.2870343.

[22] TING, F. F., Y. J. TAN, K. S. SIM. Convolutional neural network improvement for breast cancer classification. *Expert Systems with Applications*. 2019, vol. 120, pp. 103-–115. DOI: 10.1016/j.eswa.2018.11.008.

[23] SIMONYAN, K., A. ZISSERMAN. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv e-prints*. 2014, pp. arXiv:1409.1556. DOI: 10.48550/arXiv.1409.1556.

[24] TAN, M., Q. V. LE. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *36th International Conference on Machine Learning*. 2019, pp. 6105–6114.

[25] LEE, R., et al. A curated mammography data set for use in computer-aided detection and diagnosis research. *Sci Data*. 2017, vol. 4. DOI: 10.1038/sdata.2017.177.

[26] SUCKLING, J., et al. Mammographic Image Analysis Society (MIAS) database v1.21. *Apollo - University of Cambridge Repository*. 2015. DOI: 10.17863/CAM.105113.

[27] KHAMPARIA, A., et al. Diagnosis of breast cancer based on modern mammography using hybrid transfer learning. *Multidimensional Systems and Signal Processing*. 2021, vol. 32, pp. 747–765. DOI: 10.1007/s11045-020-00756-7.

[28] SHALLU, R. MEHRA. Breast cancer histology images classification: Training from scratch or transfer learning?. *ICT Express*. 2018, vol. 4, iss. 4, pp. 247–254. DOI: 10.1016/j.icte.2018.10.007.

[29] HAMEED, Z., S. ZAHIA, B. GARCIA-ZAPIRAIN, J. J. AGUIRRE, A. M. VANEGAS. Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models. *Sensors*. 2020, vol. 20, no. 16. DOI: 10.3390/s20164373.

[30] LI, H., S. ZHUANG, D.-A. LI, J. ZHAO, Y. MA. Benign and malignant classification of mammogram images based on deep learning. *Biomedical Signal Processing and Control*. 2019, vol. 51, pp. 347–354. DOI: 10.1016/j.bspc.2019.02.017.

[31] KHAN, H. N., A. R. SHAHID, B. RAZA, A. H. DAR, H. ALQUHAYZ. Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network. *IEEE Access*. 2019, vol. 7, pp. 165724–165733. DOI: 10.1109/ACCESS.2019.2953318.

[32] JAFARI, Z., E. KARAMI. Breast Cancer Detection in Mammography Images: A CNN-Based Approach with Feature Selection. *Information*. 2023, vol. 14, no. 7. DOI: 10.3390/info14070410.

A Robust Approach for Breast Cancer Classification from DICOM Images

Thanh-Nghia Nguyen
Department of Industrial Electronics and Biomedical Engineering, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam
nghiant@hcmute.edu.vn

Thanh-Tam Nguyen
Department of Industrial Electronics and Biomedical Engineering, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam
tamnt.ncs@hcmute.edu.vn

Thanh-Hai Nguyen
Department of Industrial Electronics and Biomedical Engineering, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam
nthai@hcmute.edu.vn (corresponding author)

Ba-Viet Ngo
Department of Industrial Electronics and Biomedical Engineering, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam
vietnb@hcmute.edu.vn

Received: 13 March 2025 | Revised: 3 April 2025 | Accepted: 9 April 2025

 DOI: https://doi.org/10.48084/etasr.10931

ABSTRACT

The number of breast cancer patients is rapidly increasing worldwide, with Asia currently accounting for 45% of global breast cancer cases. In addition, the number of breast cancer cases is expected to increase by 21.0%, and the mortality rate is projected to increase by 27.8% during the 2020-2030 period. This paper proposes a method for classifying breast cancer from Digital Imaging and Communications in Medicine (DICOM) images. In particular, an image segmentation technique was developed for extracting the breast region from DICOM images of varying sizes. The extracted images were then enhanced using multiple augmentation techniques to improve classification performance. Finally, a deep learning network was applied to classify breast cancer from the processed DICOM images. The VinDr-Mammo dataset was used to evaluate the effectiveness of the proposed method, and the experimental results showed an accuracy of 81.45%, demonstrating that the proposed approach is highly suitable for breast cancer detection and classification.

Keywords-breast cancer classification; deep learning; DICOM image preprocessing; image augmentation; ResNet50

I. INTRODUCTION

Currently, the number of Breast Cancer (BC) patients is rapidly increasing worldwide each year. According to statistics from the Global Cancer Organization (GLOBOCAN), in 2022, more than 2.3 million women were diagnosed with BC, with almost 670,000 deaths [1]. In Asia, the region accounts for 45% of global BC cases, and the number of cases is projected to increase by 21.0%, while the mortality rate is expected to increase by 27.8% during the 2020-2030 period [2]. In Vietnam alone, approximately 21,555 new cases and 9,345 deaths are recorded annually, accounting for 25.8% of all cancers in women [3]. BC is currently the second most common cancer among women worldwide. Early classification and treatment can significantly improve outcomes. Studies in Asian countries have shown a 5-year survival rate ranging from 56.5% to 86.7%.

In Vietnam, patients with BC diagnosed at an early stage can have a five-year survival rate of up to 90%. Moreover, a recent study on women under 35 years of age revealed that the overall 10-year survival rate for early-stage cases exceeds 80% [3]. In practice, physicians face challenges in classifying and

diagnosing BC due to difficulties in interpreting medical images (such as identifying tumors and calcifications). Furthermore, the need to analyze a large number of breast images daily can impact diagnostic accuracy. Therefore, developing a supportive system for BC diagnosis is essential. Such a system would help physicians analyze and interpret mammographic images more effectively, enabling timely and appropriate treatment decisions.

With the advancement of Artificial Intelligence (AI), Deep Learning (DL) techniques have been used effectively in BC classification, enabling early diagnosis and, thus, increasing patient survival rates [4-6]. Several recent studies have employed DL methods to classify BC using different imaging approaches. In [7], a VGG16 network model was combined with transfer learning to extract features from the BreakHis histopathological image dataset, achieving an accuracy of 89%. In [8], a model based on k-means, GMM, and CNN, utilizing Region Of Interest (ROI) for feature extraction, achieved an accuracy of 95.8%. The study in [9] focused on Lloyd's algorithm for clustering combined with CNN for classification, achieving an accuracy of 96%. These studies demonstrate the effectiveness of DL applications in BC classification.

TABLE I. A SUMMARY OF PREPROCESSING AND DEEP LEARNING METHODS FOR BC CLASSIFICATION

Ref.	Dataset (Output classes)	Preprocessing method	Model	Result
[10]	The NYU BC screening dataset V1.0 with over 1,000,000 images (benign and malignant)	Image normalization, data augmentation.	Custom ResNet-based CNN	AUC = 0.895
[11]	Assessment on the EDA dataset with 3,002 mammographic images (benign and malignant)	Removal of low-variance features, univariate feature selection, recursive feature elimination.	CNNI-BCC	High efficiency
[12]	Breast ultrasound images, MIAS, Mini-DDSM,	Wiener filtering, total variation filtering, image segmentation.	ViT, U-KAN	ViT accuracy: 99.3%, U-KAN: 93.3% accuracy
[13]	3,002 digital mammography images (benign and malignant)	CNN deep feature extraction, feature fusion.	CNN-ELM hybrid	Improved classification accuracy

Table I provides a comparative analysis of some research studies focused on BC classification using various Machine Learning (ML) and DL approaches. Each study employed different preprocessing techniques, such as image normalization, feature selection, filtering, and contrast enhancement, to improve data quality prior to classification. The classification methods range from traditional ML models [11] to advanced DL models (e.g., CNN, Vision Transformer, U-Net, and Extreme Learning Machine) [10, 12, 13]. The datasets used include mammographic images from large-scale screening examinations (eg, more than 200,000 mammograms in [10]), as well as multiple public datasets (BreastDM, MIAS, BUSI, BreakHis, DDSM) [12]. The classification outputs focus mainly on distinguishing benign and malignant tumors, with some studies extending to subtype classification [12]. The results indicate that DL models generally outperform traditional ML approaches, with CNN-based and transformer models achieving the highest accuracies (up to 99.3% in some cases). However, generalizability and computational efficiency remain challenging, highlighting the need for further DL systems to enhance the accuracy of BC classification. This comparative study underscores the rapid advancements in AI-driven BC classification and the potential of DL to improve diagnostic accuracy and support radiologists in clinical settings.

With the above statistics, BC classification plays a crucial role in current research. This paper proposes a robust method for classifying BC based on DICOM images. In particular, a DICOM image preprocessing method is applied to enhance image quality and improve classification performance. A DL network is then utilized to classify the DICOM images after preprocessing, as well as distinguishing different types of BC. The research results aim to help physicians diagnose early BC, enabling better treatment for patients. The main contributions of this study involve:

- Developing a method for extracting breast regions from DICOM images of various sizes captured by different imaging devices,
- Building a DL system for BC classification using DICOM images,
- Identifying multiple types of BC images with high accuracy.

II. MATERIAL AND METHOD

To classify different types of BC, images need to be processed and fed into a classification network. This section presents a proposed system for BC classification and the core theoretical concepts utilized.

A. Proposed Framework for BC Classification

The proposed system for BC classification, shown in Figure 1, consists of the following blocks: BC image data, image preprocessing, DL network for classification, and classification results. In particular, BC images are a DICOM dataset of breast images, which are automatically preprocessed by cropping and resizing. Data augmentation is applied to increase the number of images in the four different classes of the dataset for balancing their number of images. The training and testing blocks are designed for training and classifying different BC types. Finally, the classification results provide an evaluation of the proposed system.

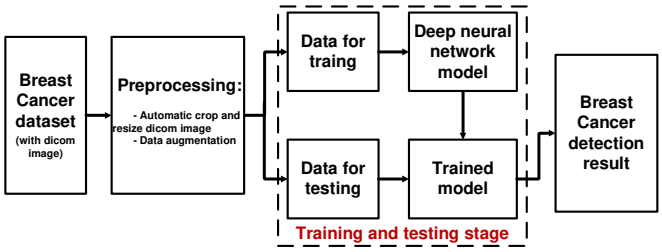


Fig. 1. Block diagram of the proposed method for BC classification.

B. Breast Cancer Dataset

This study used the VinDr-Mammo image dataset [14], which contains BC images in DICOM format, as described in Table II. DICOM images not only provide high-quality medical imaging but also store important patient information and technical parameters of the imaging device, such as patient name, age, imaging date, imaging machine type, and specific technical details. This direct connection between medical images and patient information enhances the accuracy and efficiency of diagnosis and treatment. Each patient underwent imaging in multiple views: L-CC (Left Craniocaudal), L-MLO (Left Mediolateral Oblique), R-CC (Right Craniocaudal), and R-MLO (Right Mediolateral Oblique). An example illustration of a breast cancer patient includes images from L-CC, L-MLO, R-CC, and R-MLO views.

TABLE II. BREAST CANCER DATASET [14]

Class	Label	BI-RADS	Total
1	Mass	3, 4, and 5	1113
2	Calcification	3, 4, and 5	212
3	Architectural and Asymmetry	3, 4, and 5	390
4	Normal	1 and 2	18232

C. DICOM Image Preprocessing

The VinDr-Mammo dataset was collected from multiple imaging devices, resulting in inconsistent image sizes. The dataset includes images captured from four different mammography machines: Mammomat Inspiration (image size: 3518×2800), Planmed Nuance (image size: 2812×2012), Giotto Image 3DL (image size: 3580×2812), and Giotto Class (image sizes: 3580×2531 and 3580×2543). The breast region in the images is often positioned to one side and occupies a relatively small area compared to the entire image frame. Therefore, preprocessing breast cancer images is essential. The preprocessing steps used in this study are described below.

D. Otsu and Contours Methods for Classification of Image Boundaries

The Otsu method is used to automatically determine a threshold to distinguish between the background and the object in an image. After loading the image and converting it into a NumPy array for easier processing, if the image format is not the type of uint16, it is converted to it. Therefore, the dataset containing these images with different resolutions requires standardization to be in 16-bit format. For processing these images, the Otsu method is applied to construct the image histogram, which represents the number of pixels for each grayscale value from 0 to 65,535. For determining each threshold value, the Otsu method can calculate the proportion of pixels belonging to the background and object classes. Then, the probability of each grayscale level $p(i)$ is determined, which divides the number of pixels at that grayscale level by the total number of pixels. Thus, the overall mean μ_T of the whole image is calculated using the following formula:

$$\mu_T = \sum_{i=0}^{L-1} i * p(i) \quad (1)$$

where L is the total number of grayscale levels.

The mean of the background and object classes for each threshold k are calculated as follows:

$$\mu_B(k) = \frac{\sum_{i=0}^{k} i*p(i)}{\omega_B(k)} \quad (2)$$

$$\mu_F(k) = \frac{\sum_{i=k+1}^{L-1} i*p(i)}{\omega_F(k)} \quad (3)$$

where $\omega_B(k)$ and $\omega_F(k)$ are the weights of the background and object classes, respectively. Thus, the between-class variance $\sigma_B^2(k)$ is calculated as follows:

$$\sigma_B^2(k) = \omega_F(k) * \omega_B(k)\big(\mu_B(k) - \mu_F(k)\big)^2 \quad (4)$$

Comparing all threshold values k and calculating $\sigma_B^2(k)$ for each threshold aims to determine the optimal threshold that reaches the maximum value. In addition, the output is a binary image, where pixels having values greater than the threshold are assigned the maximum value (65535), and the other pixels are assigned 0. From the obtained binary image, the contours method is used to classify the boundaries around objects in the image. Identifying contours can help to more accurately and efficiently distinguish objects from the background.

1) Creating a Bounding Box and Cropping the Image

A bounding box technique and a common image annotation method were used to accurately crop the identified object. The bounding box creates a rectangular region around the target object. After detecting the object's shape using the contour method, a bounding box is generated for the largest contour in the image. The dimensions of the bounding box correspond to the height and width of the contour. The coordinates for cropping the image at the correct position can then be calculated, ensuring the mammogram is fully captured without distortion.

2) Image Resizing Method

To ensure a uniform output size after cropping, the image is resized to 512×512 maintaining the original aspect ratio. If only one dimension is adjusted without preserving the aspect ratio, the image will be distorted. To avoid this distortion, it is crucial to calculate using a consistent scaling factor to both dimensions (width and height). The resizing process maintaining the original aspect ratio follows these steps:

- To calculate the scaling ratio for each height or width, the target size can be divided by using the original image size as follows:

$$h\_ratio = \frac{target\ height}{original\ height} \quad (5)$$

$$w\_ratio = \frac{target\ width}{original\ width} \quad (6)$$

- To prevent the image from being stretched or compressed unevenly, the smallest scaling ratio between the height and width should be selected. This ensures that the image is proportionally resized and no dimension exceeds the target size:

$$scale\_factor = (h\_ratio, w\_ratio) \quad (7)$$

- Based on the selected $scale\_factor$, the new image dimensions can be calculated to adjust proportionally both the height and width according to the following chosen ratio:

$$n\_height = o\_height * scale\_factor \quad (8)$$

$$n\_width = o\_width * scale\_factor \quad (9)$$

where o_height and n_height are the original and new heights of the image, and o_width and n_width are the original and new widths of the DICOM image, respectively.

As mentioned above, the dataset contains images of various sizes from different mammography machines. Using a standard cropping and resizing method, some images may have white padding after resizing and may not accurately capture the breast region. With the proposed cropping method, the process is automated, ensuring that the cropped images accurately match the desired breast region without unnecessary background.

3) Image Augmentation

Image augmentation is a popular technique to enhance the diversity and balance of training data. Using balanced datasets to train DL networks results in higher classification performance. Image augmentation can generate more images from the original ones without collecting additional images. In practice, common augmentation techniques include rotation, flipping, cropping, and others. Applying augmentation techniques can make the model more generalized, reduce overfitting, as well as improve performance on test data.

E. Deep Learning for Breast Cancer Classification

This study used the ResNet50 DL model to classify BC. ResNet50 consists of 50 deep layers, with its main components being residual blocks, where each block contains convolutional layers, batch normalization layers, and the ReLU activation function. The network structure is divided into five main stages, each consisting of multiple convolutional and pooling layers, enabling the model to learn complex features from the input data. The stages are described as follows:

- Conv1: The first convolutional layer has a 7×7 filter with a stride of 2 to reduce the size of the input image and capture general features. Following this, a 3×3 max pooling layer is applied to further decrease the data dimensions.
- Conv2-Conv5: These stages consist of multiple bottleneck residual blocks, each containing three consecutive convolutional layers with filter sizes of 1×1, 3×3, and 1×1, respectively. The first 1×1 layer reduces the number of channels in the input data and minimizes the number of parameters for computation. The 3×3 layer performs the main convolution operation and extracts important spatial features. The final 1×1 layer restores the number of channels to match the expected output dimensions and also ensure compatibility with the rest of the network, as described in Figure 2.

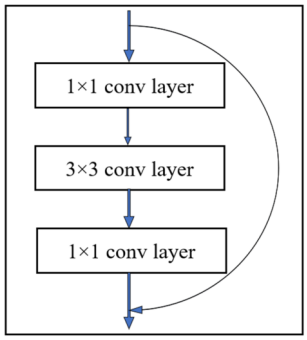


Fig. 2. The bottleneck residual block in the ResNet50 architecture.

III. RESULTS AND DISCUSSION

The proposed system's performance was evaluated using the VinDr-Mammo BC image dataset [14].

A. DICOM Image Preprocessing Results

The VinDr-Mammo BC image dataset contains numerous DICOM images with varying sizes and diverse breast region positions. Therefore, the DICOM images need to be preprocessed to extract the most distinct breast regions to train the deep learning network for BC classification. BC images were segmented using the Otsu method combined with thresholding. Next, the contours method was applied to identify objects within the images. Finally, a bounding box was generated around the breast region of each image, and the extracted breast region was cropped for classification. Figure 3 illustrates the segmentation process using Otsu, object boundary detection with contours, and the identification of the breast region using a bounding box.

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

Fig. 3. DICOM image preprocessing results: (a) original image, (b) segmentation process using Otsu, (c) boundary with object contours, (d) bounding box.

Figures 4 and 5 present the results of image cropping and resizing to 512×512. Figure 6 shows cases where the breast parts in an image are positioned on the left and right, while Figure 5 represents cases of the breasts positioned diagonally at the upper left and upper right. The results indicate that in all these cases, the cropping and resizing process effectively preserves the breast regions, ensuring high-quality outputs.

In addition, the dataset includes special cases where images have black-and-white backgrounds, as shown in Figure 6. In these cases, the image cropping algorithm still produces good results. The preprocessing results show that all images, regardless of breast position and size, can be accurately cropped to extract the necessary breast region and resized to 512×512.

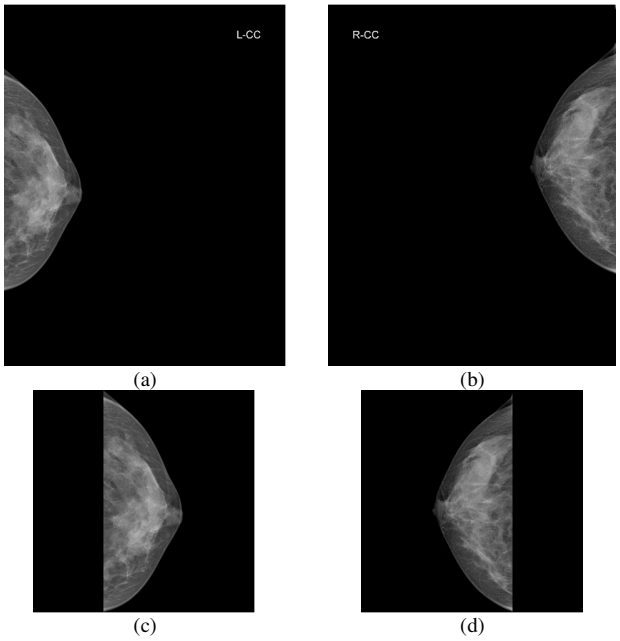


Fig. 4. Representation of the small-sized breast cancer images located on the left and right sides: (a) and (b) Original images, (c) and (d) images after cropping and resizing to 512×512.

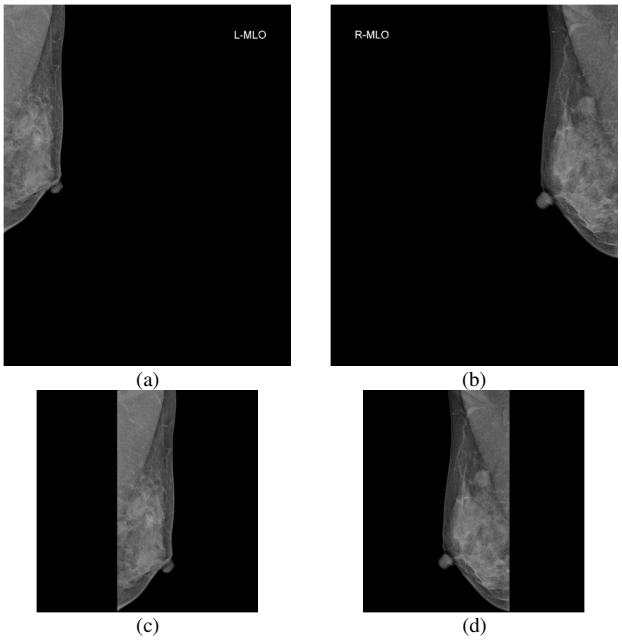


Fig. 5. Representation of the small-sized breast cancer image positioned diagonally in the top left and top right corners: (a) - (b) Original images, (c) - (d) images after cropping and resizing to 512×512.

B. Breast Cancer Classification result

The dataset has a significant imbalance in the number of images across classes. In particular, the number of images in classes 2 and 4 is 212 and 18,232, respectively. Therefore, data augmentation is essential to improve classification performance. Image augmentation involved vertical and horizontal flipping and image rotation within the range of ±50°.

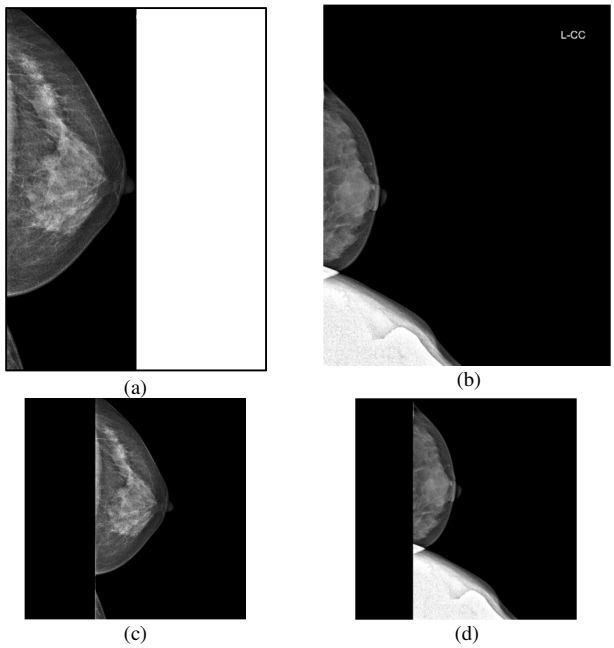


Fig. 6. Representation of the breast cancer images with some special cases: (a) and (b) original images, (c) and (d) images after cropping and resizing to 512×512.

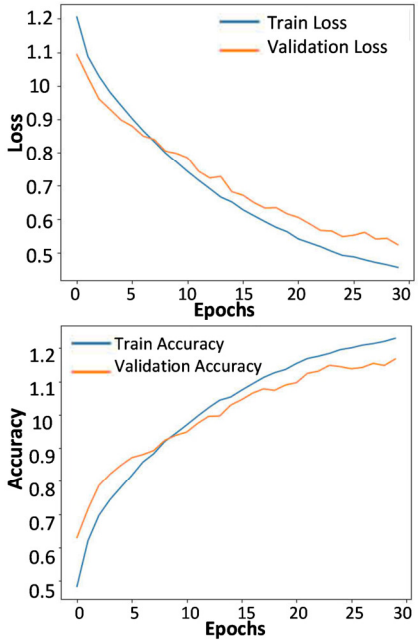


Fig. 7. Training results for Case 3 (10,000 images per class).

Four experimental scenarios were tested: In Case 1, each class contains 2,000 images, in Case 2, each class contains 5,000 images, in Case 3, each class contains 10,000 images, and in Case 4, each class contains 10,000 images, except for class 4, which retains 18,232 images. For class 4, a random selection was applied to match the number of images in the other classes. The total dataset was split into 80% for training, 10% for validation, and 10% for testing the model.

The ResNet50 model was used to classify BC images. The images were resized to be 512×512 and trained the network for 30 epochs. Figure 7 presents the training results for Case 3, where the number of processed images per class is 10,000. The results show that the loss continuously decreased and the accuracy improved significantly as the number of epochs increased. Figure 8 presents the confusion matrix for Case 3, in which the classification performance for the Calcification category is excellent. However, the classification performance for the Mass category is lower.

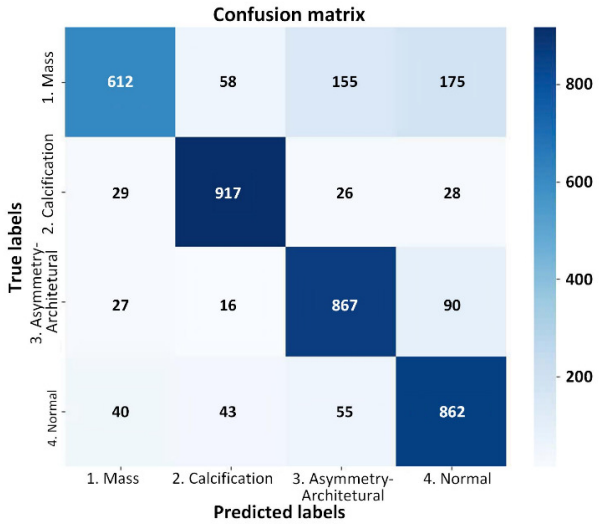


Fig. 8. Confusion matrix for Case 3 (10,000 images in each class).

Based on the confusion matrix results, evaluation metrics such as Accuracy (Acc), Precision (Pre), Recall (Rec), and F1-score (F1s) were calculated to evaluate the classification performance across different classes [15]. Figure 8 presents the confusion matrix results for Case 3, where it can be seen that the classification performance across all classes is relatively good. In particular, the Calcification category had the highest classification accuracy of 92%, while the Mass category had an accuracy of 61%. Table III summarizes the image classification results for Case 3.

TABLE III. RESULTS FOR BREAST CANCER CLASSIFICATION

Class	Label	Acc	Pre	Rec	F1s
1	Mass	0.61	0.86	0.61	0.72
2	Calcification	0.92	0.89	0.92	0.90
3	Architectural-Asymmetry	0.87	0.79	0.87	0.82
4	Normal	0.86	0.75	0.86	0.80

Table VI presents the comparative performance results of the four cases. In Case 1, where the number of images per class was augmented to 2,000, classification performance was the lowest, with results for Acc, Pre, Rec, and F1s of 72.92%, 75.64%, 72.92%, and 70.66%, respectively. As the number of augmented images increased, the classification performance improved. In practice, Acc reached 77.83% when the number of images increased to 5,000 and 81.45% when it increased to 10,000. However, in Case 4, since the number of images in the Normal class was 18,232, the classification performance decreased when increasing the number of images in the Mass, Calcification, and Architectural and Asymmetry classes to 10,000. In this case, the results were Acc=81.65%, Pre=84.14%, Rec=78.41%, and F1s=79.95%. These results indicate that the classification performance across different classes was inconsistent, leading to performance discrepancies between classes.

TABLE IV. PERFORMANCE COMPARISON FOR THE FOUR CASES

	Case 1	Case 2	Case 3	Case 4
Acc	72.92	77.83	81.45	81.65
Pre	75.64	77.80	82.09	84.14
Rec	72.92	77.84	81.45	78.41
F1s	70.66	77.82	81.07	79.95

DL has been widely applied for classification in various domains, including agriculture, healthcare, aviation, and industry. In the medical field, DL is increasingly utilized for early diagnosis, disease classification, and other applications. For BC classification, the ResNet50 deep learning model has been employed with promising results [16-18]. In particular, in [16] ResNet50 was applied for BC classification on the BreakHis open dataset, achieving a high accuracy of 92.24%, but this system was limited to distinguishing between benign and malignant tumor types. In [17], ResNet50 used a self-collected breast cancer dataset consisting of 2,088 images classified into Healthy and Sick categories, achieving an accuracy of more than 80% [17]. This study applied ResNet50 to classify four different subtypes of BC, achieving an accuracy of 81.45%.

The VinDr-Mammo dataset consists of BC images collected from Vietnamese patients. Since this is a relatively new dataset, there are limited studies on it. Research on the VinDr-Mammo could provide valuable support for BC diagnosis in Vietnam. Moreover, this study serves as a foundation for developing more advanced algorithms to enhance BC detection and classification in the country. This study proposes a combination of DICOM image preprocessing and the ResNet DL model to classify different BC subtypes using this dataset. The results indicate that increasing the number of images per class to 10,000 improves classification performance. However, when the Normal class remains at 18,232 images, the classification performance across different classes is still suboptimal. Future research will aim to expand the dataset and enhance deep learning models to further improve classification performance. This study could serve as a basis for scientists to develop algorithms to improve BC detection and classification.

IV. CONCLUSION

Early and accurate diagnosis of BC can significantly improve patient recovery outcomes. This study combined an automatic image segmentation method with DL for BC classification in DICOM images. The experimental results demonstrated that the proposed method is well-suited for this purpose. In particular, when data augmentation was applied to increase the number of images per class to 10,000, classification performance improved significantly, achieving

Acc=81.45%, Pre=82.09%, Rec=81.45%, and F1s=81.07%. However, due to the imbalance in the number of images across different classes, the classification performance for certain classes remained lower, requiring further augmentation of the underrepresented classes. This research on BC detection and classification using the VinDr-Mammo dataset can serve as a foundation for scientists to develop further studies using this dataset. In the future, larger and more diverse datasets will be explored and integrated to create a more balanced dataset for improving BC detection and classification.

ACKNOWLEDGMENT

This work is supported by Ho Chi Minh City University of Technology and Education (HCMUTE) under Grant No. T2024-133.

REFERENCES

[1] J. Kim *et al.*, "Global patterns and trends in breast cancer incidence and mortality across 185 countries," *Nature Medicine*, vol. 31, no. 4, pp. 1154–1162, Apr. 2025, https://doi.org/10.1038/s41591-025-03502-3.

[2] A. Elhusseiny, "Women's cancer is getting worse in Asia Pacific," *World Economic Forum*, Oct. 11, 2023. https://www.weforum.org/stories/2023/10/womens-cancer-is-getting-worse-in-asia-pacific-heres-what-to-do/.

[3] VnExpress, "Novartis sponsors expert talkshow to raise breast cancer awareness - VnExpress International," *VnExpress International*. https://e.vnexpress.net/news/business/novartis-sponsors-expert-talkshow-to-raise-breast-cancer-awareness-4544893.html.

[4] L. C. V Priya, V. G. Biju, B. R. Vinod, and S. Ramachandran, "Deep learning approaches for breast cancer detection in histopathology images: A review," *Cancer Biomarkers*, vol. 40, no. 1, pp. 1–25, May 2024, https://doi.org/10.3233/CBM-230251.

[5] B. Asadi and Q. Memon, "Efficient breast cancer detection via cascade deep learning network," *International Journal of Intelligent Networks*, vol. 4, pp. 46–52, Jan. 2023, https://doi.org/10.1016/j.ijin.2023.02.001.

[6] S. M. Shaaban, M. Nawaz, Y. Said, and M. Barr, "An Efficient Breast Cancer Segmentation System based on Deep Learning Techniques," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12415–12422, Dec. 2023, https://doi.org/10.48084/etasr.6518.

[7] D. Albashish, R. Al-Sayyed, A. Abdullah, M. H. Ryalat, and N. Ahmad Almansour, "Deep CNN Model based on VGG16 for Breast Cancer Classification," in *2021 International Conference on Information Technology (ICIT)*, Amman, Jordan, Jul. 2021, pp. 805–810, https://doi.org/10.1109/ICIT52682.2021.9491631.

[8] S. Shamy and J. Dheeba, "A research on detection and classification of breast cancer using k-means GMM & CNN algorithms," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6S, pp. 501–505, 2019.

[9] V. S. Vijayan and P. L. Lekshmy, "Deep learning based prediction of breast cancer in histopathological images," *International Journal of Engineering Research & Technology*, vol. 8, no. 07, pp. 148–152, 2019.

[10] N. Wu *et al.*, "Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1184–1194, Apr. 2020, https://doi.org/10.1109/TMI.2019.2945514.

[11] A. Khalid *et al.*, "Breast Cancer Detection and Prevention Using Machine Learning," *Diagnostics*, vol. 13, no. 19, Jan. 2023, Art. no. 3113, https://doi.org/10.3390/diagnostics13193113.

[12] Z. Zhu, Y. Sun, and B. Honarvar Shakibaei Asli, "Early Breast Cancer Detection Using Artificial Intelligence Techniques Based on Advanced Image Processing Tools," *Electronics*, vol. 13, no. 17, Jan. 2024, Art. no. 3575, https://doi.org/10.3390/electronics13173575.

[13] Z. Wang *et al.*, "Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features," *IEEE Access*, vol. 7, pp. 105146–105158, 2019, https://doi.org/10.1109/ACCESS.2019.2892795.

[14] H. H. Pham, H. Nguyen Trung, and H. Q. Nguyen, "VinDr-Mammo: A large-scale benchmark dataset for computer-aided detection and diagnosis in full-field digital mammography." PhysioNet, https://doi.org/10.13026/BR2V-7517.

[15] T. N. Nguyen, T. H. Nguyen, M. H. Nguyen, and S. Livatino, "Wavelet-Based Kernel Construction for Heart Disease Classification," *AEEE Advances in Electrical and Electronic Engineering*, vol. 17, no. 3, pp. 306–319, Sep. 2019, https://doi.org/10.15598/aeee.v17i3.3270.

[16] N. Behar and M. Shrivastava, "ResNet50-Based Effective Model for Breast Cancer Classification Using Histopathology Images," *CMES - Computer Modeling in Engineering and Sciences*, vol. 130, no. 2, pp. 823–839, Dec. 2021, https://doi.org/10.32604/cmes.2022.017030.

[17] S. Malathi, "Breast Cancer Detection With Resnet50, Inception V3, And Xception Architecture.," *Journal of Pharmaceutical Negative Results*, vol. 14, no. 4, pp. 60-68, 2023.

[18] E. Al. T Sunil Kumar, "Breast Cancer Classification and Predicting Class Labels Using ResNet50," *Journal of Electrical Systems*, vol. 19, no. 4, pp. 270–278, Jan. 2024, https://doi.org/10.52783/jes.638.

Segmentation and Color ROI Extraction from Breast Imaging Datasets for Cancer Classification

Thanh Tam Nguyen
Department of Industrial Electronics and Biomedical Engineering, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam
tamnt.ncs@hcmute.edu.vn

Thanh Hai Nguyen
Department of Industrial Electronics and Biomedical Engineering, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Vietnam
nthai@hcmute.edu.vn (corresponding author)

Received: 11 May 2025 | Revised: 19 June 2025 | Accepted: 5 July 2025

 DOI: https://doi.org/10.48084/etasr.12067

ABSTRACT

Breast cancer is a serious health concern worldwide, particularly affecting many women. Therefore, using image processing techniques to enhance mammographic images is essential for accurate and early diagnosis. This study proposes a method for segmenting and extracting the Gray Region of Interest (GROI) from mammographic images and creating a corresponding Color Region of Interest (CROI) to improve classification performance. An EfficientNet-B7 model is used to classify image sets containing CROI. To accurately evaluate the effectiveness of CROI compared to GROI, the proposed method is applied to five categories of mammography image sets before training the EfficientNet-B7 model. Specifically, an automated algorithm is introduced to determine the thresholds values for extracting the GROI. The CROI is the generated using the proposed Identification and Comparison (IaC) algorithm. The results show that the classification accuracy improves from 84.3% with GROI to 92.6% with CROI, demonstrating the effectiveness of the proposed enhancement method for breast cancer image classification.

Keywords-five mammography categories; IaC algorithm for CROI; segmentation algorithm for GROI; breast cancer classification

I. INTRODUCTION

Breast cancer is a malignant disease that develops from breast cells. Although it is most common in women, men can also develop it. Breast lesions are often classified into many different types, including mass, suspicious calcification, asymmetry (focal or global), architectural distortion, skin thickening, skin retraction, and nipple retraction, among others [1-4]. Each type of breast lesion in these images can be a sign of breast cancer. Therefore, determining the correct type of lesion is an important part of the diagnosis process and can assist doctors accurately assess cancer risk. Early breast cancer detection increases the chances of successful treatment with less invasive methods. Moreover, early detection based on mammography images can significantly improve the survival rate of breast cancer patients. According to statistics, women diagnosed in the first stage of breast cancer have a 5-year survival rate of up to 99%, which is much higher than for those diagnosed in later stages. Besides, analyzing mammography images is a safe, non-invasive technique suitable for most women aged 40 and over [5].

In practice, many different image segmentation methods exist to identify objects known as Regions of Interest (ROIs). These methods include region-based methods, boundary-based methods, atlas-based methods, and model-based methods, as well as deep learning in medical images [6, 7]. These methods can be used to remove the background and noise and keep only the object to be evaluated, called a tumor, in cancer images [8, 9].

One of the most popular methods is the Otsu method, which is widely used in medical images such as X-ray and CT scan images [10]. Specifically, the Otsu method was applied in conjunction with the Hounsfield unit to determine the threshold value needed to convert CT scan images into binary images. Using this method, the highest segmentation result was 77.43%. In this paper, an optimal Otsu threshold is used to extract the ROI from gray images with a black background from a mammography image set containing four lesion

categories. After processing the images to produce the Gray Region of Interest (GROI) with enhanced contrast and brightness, the resulting images are used for training neural networks. Notably, combining image enhancement techniques with Artificial Intelligence (AI) models has shown improved classification performance [11, 12]. Adaptive contrast enhancement algorithms are often applied to achieve uniform contrast, contributing to better outcomes when using AI models. To further improve image quality, Adaptive Gamma Correction Weighted Distribution (AGCWD) has been employed to optimize contrast and brightness, while Recursively Separated and Weighted Histogram Equalization (RSWHE) preserves brightness and enhance contrast [13].

Images containing GROIs have been used in various domains with different techniques. For example, the watershed algorithm was applied for over-segmentation using the luminance and saturation characteristics of color images to enable automatic ROI extraction [14]. In another study, focused on printed fabric patterns, the mean shift algorithm combined with color measurements enabled consistent and reliable detection of multiple color patterns [15]. In our work, mammography images with Color Regions of Interest (CROIs) are used for breast cancer classification. The EfficientNet-B7 model is employed, demonstrating its effectiveness when learning from CROI-enhanced image sets.

Deep-learning models require large, balanced datasets and data augmentation is a common approach to expand the data. Typical techniques include geometric transformation, pixel level augmentation, pseudo-color augmentation; random erasing, and kernel filters [16]. In this study, we use rotation and flipping to increase the dataset size. With this image set enhancement and balancing, the EfficientNet-B7 classifier can achieve high classification accuracy.

AI is widely applied in the medical field, particularly for diagnosing breast cancer based on mammography images [17-19]. Recent studies have applied Convolutional Neural Networks (CNNs) for mammography classification [20-22]. For instance, AlexNet achieved 83.1 % accuracy on the DDSM dataset when classifying normal images, Invasive Breast Cancer (IBC), and Ductal Carcinoma in Situ (DCIS) [23], whereas the GoogLeNet network achieved 96.37% accuracy on a separately collected dataset [24].

With the development of deep learning networks, the classification accuracy has improved. VGG16 achieved sensitivity and specificity of 95.4 % and 98.3 %, respectively, on the CBIS dataset [25], and ResNet50 achieved 96.2 % sensitivity and 99.1 % specificity on the INbreast dataset [26]. Furthermore, DenseNet classified malignancy with 92.5% accuracy based on five categories on the Breast Cancer Histopathology Image Database dataset [27], whereas AlexNet achieved 83.4% [28].

EfficientNet began with version B0 and expanded to B1–B7. EfficientNet-B7 achieves higher accuracy compared to other deep learning models, especially on ImageNet, while being 8.4× smaller and 6.1× faster than current deep learning networks [29]. In recent researches, EfficientNet models have been combined with other algorithms in many different classification fields. In particular, Lung-EfficientNet was proposed for lung cancer classification based on CT-scan image sets, achieving an accuracy of 99.10% [30]. EfficientNet based U-Net models have been applied for the segmentation of kidney tumors on CT-scan images, producing highly accurate kidney-tumor segmentation [31, 32]. In this paper, we appropriately fine-tune EfficientNet-B7 to classify images with CROIs on the VinDr-Mammo dataset, achieving high accuracy.

This study aims to develop a method for processing breast lesion image datasets that maximizes the effectiveness of classification before applying it to a deep learning network. Classification is performed on five main types of lesions: mass, calcification, architectural distortion, asymmetry, and normal (no finding). The proposed method applies image segmentation to extract GROIs and then converts them to CROIs to enhance classification performance. Since most deep learning networks are pre-trained on three-channel image datasets such as ImageNet, converting single-channel mammograms into CROIs aligns the data format with the network's input requirements. Additionally, this study uses a relatively new set of mammography images called the VinDr-Mammo image dataset [33], collected from hospitals in Vietnam.

II. MATERIALS AND METHODS

This study assesses the effectiveness of the proposed method by applying both GROIs and CROIs of breast cancer images before training the EfficientNet-B7 model. Mammography image sets are categorized into five lesion categories: mass, calcification, architectural distortion, asymmetry, and normal. Classification using the EfficientNet-B7 model is performed in two steps: first, GROIs are extracted from the image sets, and then CROIs are converted to GROIs. Furthermore, to balance the datasets, images are increased, and image sets that are overrepresented are randomly reduced. To augment the images, two geometric transformation methods are used: rotation and flipping. With this balance, the classification accuracy using an EfficientNet-B7 will increase.

As shown in Figure 1, the following methods are applied to the mammography image sets: extracting GROI, converting GROI into CROI, augmenting the image sets for balance, and using an EfficientNet-B7 model. After balancing, the image sets are fed into the EfficientNet-B7 model, initialized with pre-trained ImageNet weights. The model is fine-tuned by modifying the final layers and adding appropriate output layers. Classification results are then compared to determine the most effective approach for classifying breast mammography images across five lesion categories. Figure 1 presents the proposed framework, which includes the VinDr-Mammo dataset [33], the blocks for extracting GROI and CROI, data augmentation, and the EfficientNet-B7 classifier to evaluate the classification of three cases (original images, GROI, and CROI) across five breast lesion categories.

A. Introduction to the VinDr-Mammo Dataset

In this research, we used the VinDr-Mammo dataset [33], which contains mammography images collected between 2018 and 2020. These image sets were stored in the Picture Archiving and Communication System (PACS) of Hanoi Medical University Hospital (HMUH) and Hospital 108

(H108). The dataset includes 5,000 mammograms scans, corresponding to 20,000 mammography images, because each scan produces four breast images including: Right Cranial-Caudal (CC), Left CC, Right Mediolateral-Oblique (MLO), and Left MLO, as illustrated in Figure 2. It is important to note that the personal information of the patients in the images is always protected, and these images only contain information related to lesions. The VinDr-Mammo dataset categorizes these images into five lesion types and eleven breast abnormalities, including mass, suspicious calcification, asymmetry, focal asymmetry, global asymmetry, architectural distortion, skin thickening, skin retraction, nipple retraction, and suspicious lymph nodes. The age of patients ranges from 20 to 86 years, with the highest concentration in the 40–45 age group. Since the dataset is imbalanced with regard to the different lesion categories, we only selected four categories of lesions and one category without lesions for this research, as described in Table I.

TABLE I. DESCRIPTION OF LESION CATEGORIES IN THE VINDR-MAMMO BREAST IMAGE DATASET

Lesion category	Number of images
Mass	1,226
Suspicious calcification	543
Focal asymmetry	269
Architectural distortion	119
Normal (no finding)	2,551

B. Balance of Mammography Image Sets

The image sets for the five lesion categories have different sizes. Therefore, transformation methods such as flipping and rotating images were applied to increase the number of images to 3,000 for each category.

In the flipping method, images are flipped horizontally and vertically to create mirror images of the originals. Assuming $pin(x, y)$ is the original image at coordinates (x, y), where $x \in [0, w-1]$ and $y \in [0, h-1]$, flipping is described as follows:

- Horizontal flip: $p_{outw}(x, y) = p_{inw}(w - x, y)$
- Vertical flip: $p_{outh}(x, y) = p_{inh}(x, h - y)$

Here, $p_{inw}(x, y)$ and $p_{inh}(x, y)$ are the pixels at (x, y) in the input image, whereas $p_{outw}(x, y)$ and $p_{outh}(x, y)$ denote the pixels of the output flipped image. The parameters w and h denote the width and height of the image in pixels, respectively.

For image rotation, images are rotated around the center by different angles to generate variations in the GROI position. In this study, images are rotated by 45°, 90°, 135°, 180°, 225°, 270°, and 315°. Assuming that an original pixel at (x, y) moves to a new position (x_r, y_r), the new coordinates are computed as:

$$x_r = r\cos(\alpha + \theta) \quad (1)$$

$$y_r = r\sin(\alpha + \theta) \quad (2)$$

where $r = \sqrt{x^2 + y^2}$, $\alpha = \arctan\left(\frac{y}{x}\right)$, and θ is the rotation angle.

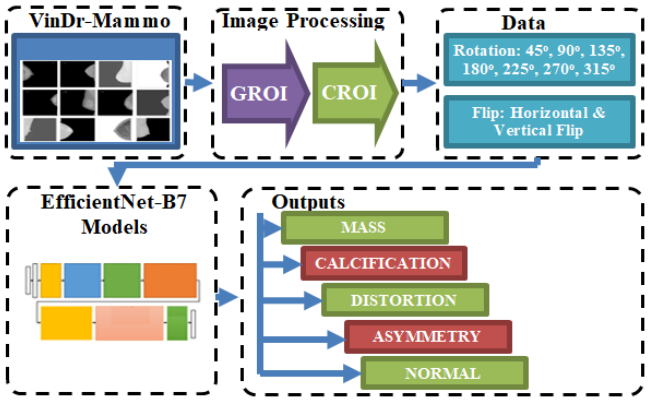


Fig. 1. Block diagram of the proposed method for breast lesion classification.

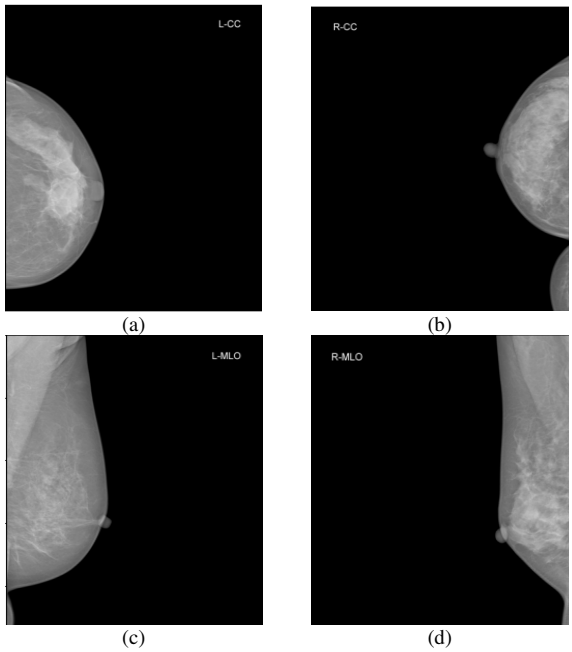


Fig. 2. A sample set of four breast images from a single mammography exam in the VinDr-Mammo dataset: (a) left CC, (b) right CC, (c) left MLO, (d) right MLO.

C. Extraction of Gray Region of Interest

Image preprocessing is an important step in classifying images using deep learning networks. In fact, applying preprocessing techniques to mammography images enables deep learning networks to achieve more accurate classifications and reduce training time. This study applies Otsu segmentation to extract images with GROI, which are then converted into images with CROI.

To extract an effective GROI without losing important image features, the Otsu segmentation is applied to determine an optimal threshold. Specifically, the images are segmented based on this threshold to produce images with GROI and a black background [8]. Assuming the gray levels in the input image $I(x, y)$ are $k = 0, 1, 2 \ldots, L-1$, a threshold k is chosen to divide the image pixels into two sets, C_0 and C_1. The optimal

threshold is found based on the variance between C_0 and C_1 using the following equation:

$$T_{opt} = \text{argmax}_T(\omega_0(\mu_0 - \mu_T)^2 + \omega_1(\mu_1 - \mu_T)^2) \quad (3)$$

where μ_0, μ_1 are the average gray levels of the sets C_0 and C_1, and μ_T is the average gray level of the image. The Algorithm 1 for determining the optimal Otsu threshold is summarized in Figure 3. For each original image, an adaptive optimal threshold using the Otsu method is determined, followed by segmentation to produce the ROI. Next, morphological operations (TopHat and BlackHat) and a cleaning mask are applied to produce the final image with GROI and a black background.

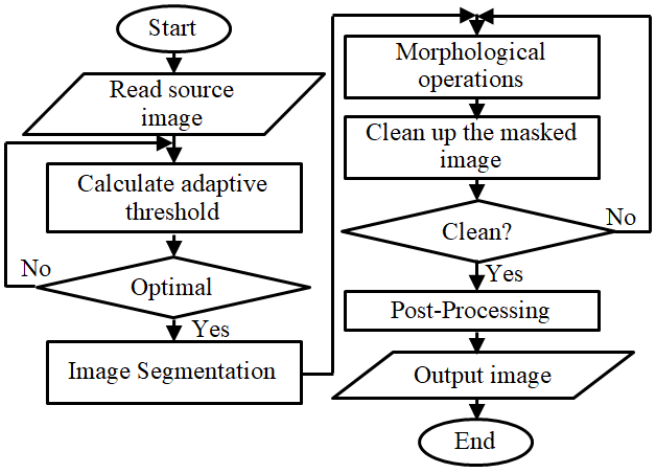


Fig. 3. Workflow of Algorithm 1 for determining the Otsu threshold to extract the GROI from mammography images.

D. Extraction of Color Region of Interest

Mammography images are grayscale images with one color channel, whereas deep learning models, such as the EfficientNet-B7 model, are pre-trained on the ImageNet dataset and can effectively work with color images that have three color channels. To align the data with these models, each GROI is converted into a CROI that highlights the tumor, dense tissues, and surrounding cells in different colors.

To create the CROI, we developed an algorithm that visualizes features that need to be identified, such as tumor areas in mammogram images, by applying a selected color gradient to highlight the lesion areas. The algorithm consists of two tasks: Identification and Comparison (IaC). In the identification task, areas in the GROI that exhibit abnormal gray levels are marked as signs of tumors or other types of lesions. The comparison task involves detecting differences between regions of a small group of points and the surrounding background, which may indicate signs of damage or calcification. For the first task, a gray level index is calculated and then converted into color levels to add to the ROI image. Meanwhile, suspicious areas such as calcifications are detected using two morphological transformations: TopHat and BlackHat. The final CROI image I_{CROI} is produced using the following expression:

$$I_{CROI} = (G(x,y) + I_{TH}) - I_{BH} - Z|[SE_Z^{\wedge} \cap G(x,y)] \subseteq G(x,y) \quad (4)$$

where $G(x,y)$ is the input image, SE is the structuring element of the image, and I_{TH} and I_{BH} are determined as follows:

$$I_{TH} = G(x,y) - (G(x,y) \ominus SE) \oplus SE \quad (5)$$

$$I_{BH} = (G(x,y) \oplus SE) \ominus SE - G(x,y) \quad (6)$$

Based on (4), the algorithm for converting a GROI into a CROI is described as follows:

```
Algorithm 2. Converting GROI into CROI
1. START
2  READ source image G(x,y)
3. Generate dilated image I_D from the
source image
4. Calculate TopHat image I_TH
5. Calculate BlackHat image I_BH
6. Generate ColorROI image I_CROI
7. OUTPUT I_CROI
8. END
```

E. CNN Model for Lesion Classification

This research applies an EfficientNet-B7 model to evaluate the classification performance for three cases: original images, images with GROI, and images with CROI. Specifically, GROI is extracted from the image sets and converted into images with CROI, which are then fed into the EfficientNet-B7 model for classification. The proposed model employs EfficientNet-B7, which has approximately 66 million parameters, as illustrated in Figure 4. This network is highly suitable for color images and produces high classification performance.

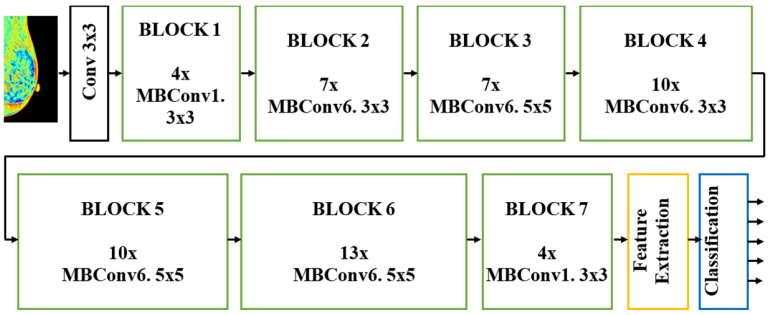


Fig. 4. Architecture of the fine-tuned EfficientNet-B7 for classifying five categories of mammography image sets.

Furthermore, EfficientNet-B7 is designed not only to increase classification accuracy, but also to improve model performance by reducing the number of parameters, as well as the computational load, measured as Floating Point Operations per Second (FLOPS), compared to other models. EfficientNet-B7 uses a new, complex expansion method with an aggregation factor that uniformly scales the width, depth, and resolution, enabling it to achieve better accuracy.

In this article, the EfficientNet-B7 model includes 9 blocks and a total of 59 layers. The first 7 blocks are used for separating features from GROI images that are converted into RGB 3-channel color images. These blocks use weights that were pre-trained on the ImageNet dataset. The feature extraction and classification blocks are added to perform the task of classifying the input images into five lesion categories. Moreover, training is performed on a computer with an Intel Core i9-9980 processor, 32 GB of DDR4 RAM, and a Tesla P4 graphics card. During training, the network runs for 350 epochs under established stopping conditions. Furthermore, the Adam optimization algorithm is used during training with a learning rate of 0.001.

III. RESULTS AND DISCUSSION

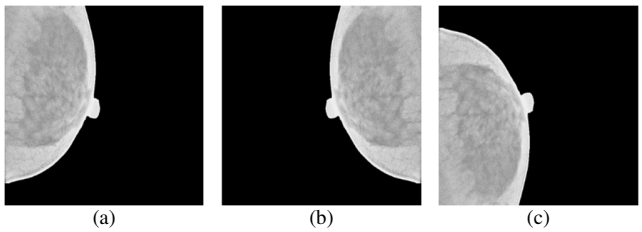
A. Augmentation of Image Sets

The original VinDr-Mammo dataset provides grayscale Digital Imaging and Communications in Medicine (DICOM) images at a resolution of 2012×2812, which were converted to JPEG format before being processed into the next steps. After segmentation and image enhancement, it is necessary to balance the five image sets. For example, the ROI image in Figure 5(a) is flipped horizontally and vertically, as shown in Figures 5(b) and 5(c), respectively, creating two additional images. Another example of augmentation is rotating an image at different angles to create multiple images. In this study, we rotate an image at seven different angles: 45°, 90°, 135°, 180°, 225°, 270°, and 315° creating seven additional images, as shown in Figure 6.

After image preprocessing and data augmentation, the images are classified using the EfficientNet-B7 model. The original dataset was divided into training and test sets at a ratio of 0.8 to 0.2. These sets were augmented independently, as shown in Table II.

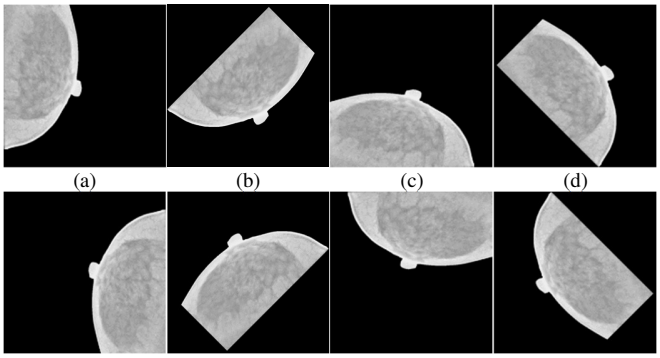
TABLE II. ORIGINAL IMAGE COUNTS, TRAIN/TEST SPLITS, AND BALANCED DATASET SIZES AFTER AUGMENTATION OR REDUCTION

Category	Original	Train	Test	Balanced train	Balanced test
Mass	1,226	981	245	2,400	600
Suspicious calcification	543	434	109	2,400	600
Focal asymmetry	269	215	54	2,400	600
Architectural distortion	119	95	24	2,400	600
Normal (no finding)	2,551	2,041	510	2,400	600
Total	4708	3766	942	12000	3000



(a) (b) (c)

Fig. 5. Representation of flipped images: (a) original GROI image, (b) horizontally flipped image, (c) vertically flipped image.



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

(e) (f) (g) (h)

Fig. 6. Representation of rotated images: (a) original GROI image, (b) image rotated by 45, (c) 90°, (d) 135°, (e) 180°, (f) 225°, (g) 270°, (h) 315°.

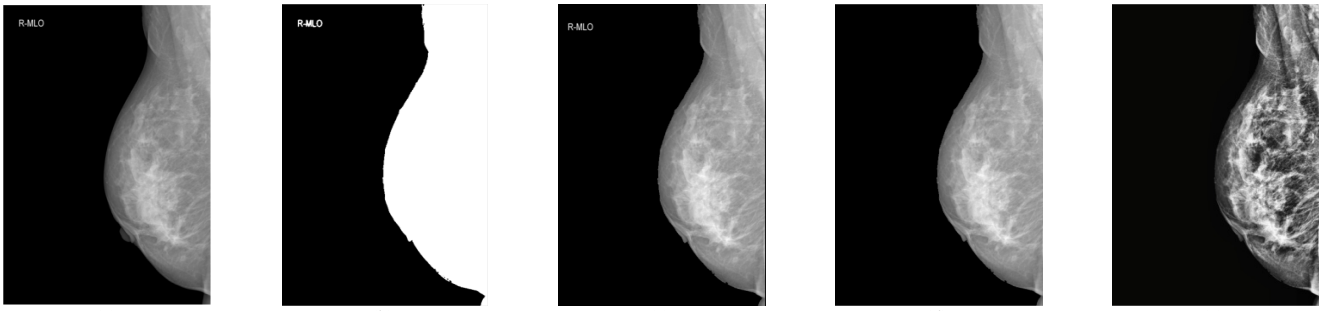
B. Results of GROI and CROI Extraction

Original breast lesion images often contain a lot of information about the patient and other components that can affect the classification process. Algorithm 1 removes these components and retains only the GROI and the black background. Figure 7 illustrates these image processing steps, whereas Figure 8 shows representative GROIs from the five lesion classes.

Subsequently, the IaC algorithm converted each GROI to a CROI, highlighting suspicious structures. Figure 9 shows the results for the five lesion categories, where the colored areas in the CROI are distinct from normal tissue regions, indicating areas at high risk of breast cancer. Thus, these CROI images can enable highly accurate classification using EfficientNet-B7.

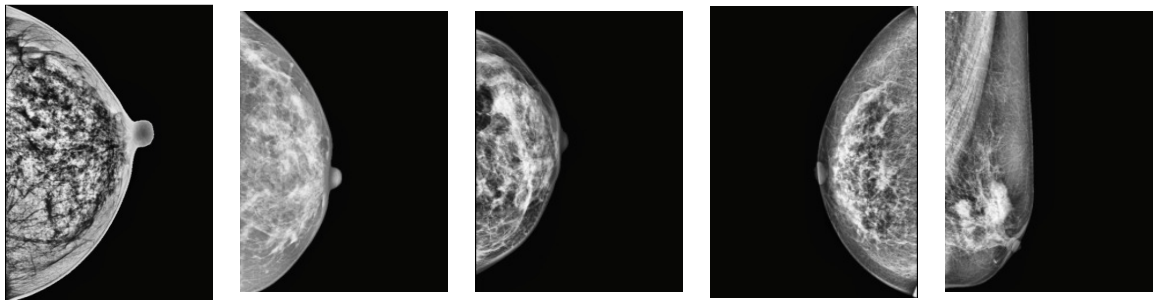
In Figure 9, images with CROI show signs of abnormalities through different color characteristics compared to cases without lesions. Specifically, red areas may indicate tumors, whereas calcifications appear as blue regions with red spots. This clear color differentiation suggests that the classification performance on images with CROI will be high using deep learning networks.

Figure 10 compares the confusion matrices for GROI and CROI inputs. The rows represent the actual labels of image sets, whereas the columns represent the predicted labels. The diagonal elements indicate the number of True Positives (TP). The results show that the classification accuracy for images with CROI (Figure 10(b)) is superior to that of images with GROI. Specifically, the average classification accuracy is 92.6% for CROI compared to 84.3% for GROI. This demonstrates the effectiveness of the EfficientNet-B7 model when applied to images with CROI.



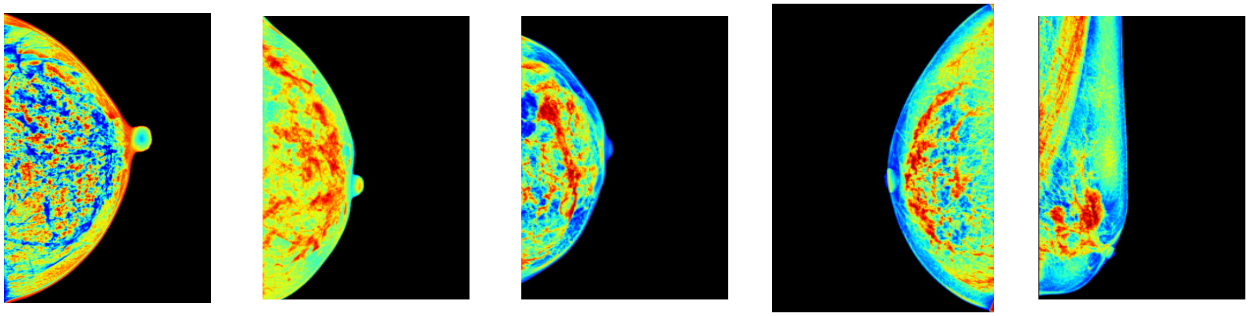
(a) (b) (c) (d) (e)

Fig. 7. Image processing workflow for GROI extraction: (a) original image, (b) masted image using optimal threshold, (c) segmented image, (d) image after morphological operations and clean up, (e) image with GROI after postprocessing.



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

Fig. 8. Representation of five images categories with GROI and black background: (a) normal, (b) focal asymmetry, (c) architectural distortion, (d) calcification, (e) mass.



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

Fig. 9. Representation of five images categories with CROI: (a) normal, (b) focal asymmetry, (c) architectural distortion, (d) calcification, (e) mass.

To verify the reliability of the results, the classification was repeated three times to obtain statistical consistency. Table III presents the results from the three iterations along with their averages. The simulation outcomes confirm that the proposed image preprocessing methods significantly increase overall accuracy, demonstrating that the EfficientNet-B7 model is highly suitable for the image dataset with CROI. Furthermore, Table IV presents the precision, recall, and F1 score for each lesion category when using CROI inputs.

C. Comparison with Related Work

In recent years, numerous studies have focused on the classification of breast images. One such study focused on classifying images with benign and malignant masses using multiple classifiers, including Support Vector Machine (SVM), and Neural Networks (NN) [12]. In particular, 320 mammograms from 80 patients were enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE) and then segmented to produce optimal threshold values before feature extraction. The classification results demonstrated a high accuracy rate of 97% using NN. These results illustrate the effectiveness of image processing before entering the classifier.

Another study applied a variety of preprocessing techniques, including image resizing, data normalization, and data augmentation, to prepare data for analysis. These techniques optimized the format and improved the model's generalizability [34]. After preprocessing, the image sets were trained and analyzed using an Ensemble Deep Convolutional Neural Network (EDCNN) model. The dataset included 943 ultrasound images and was divided into Dataset-1 (780 images) and Dataset-2 (163 images). The results demonstrated an exceptional accuracy: 87.82% for Dataset-1 and 85.69% for Dataset-2.

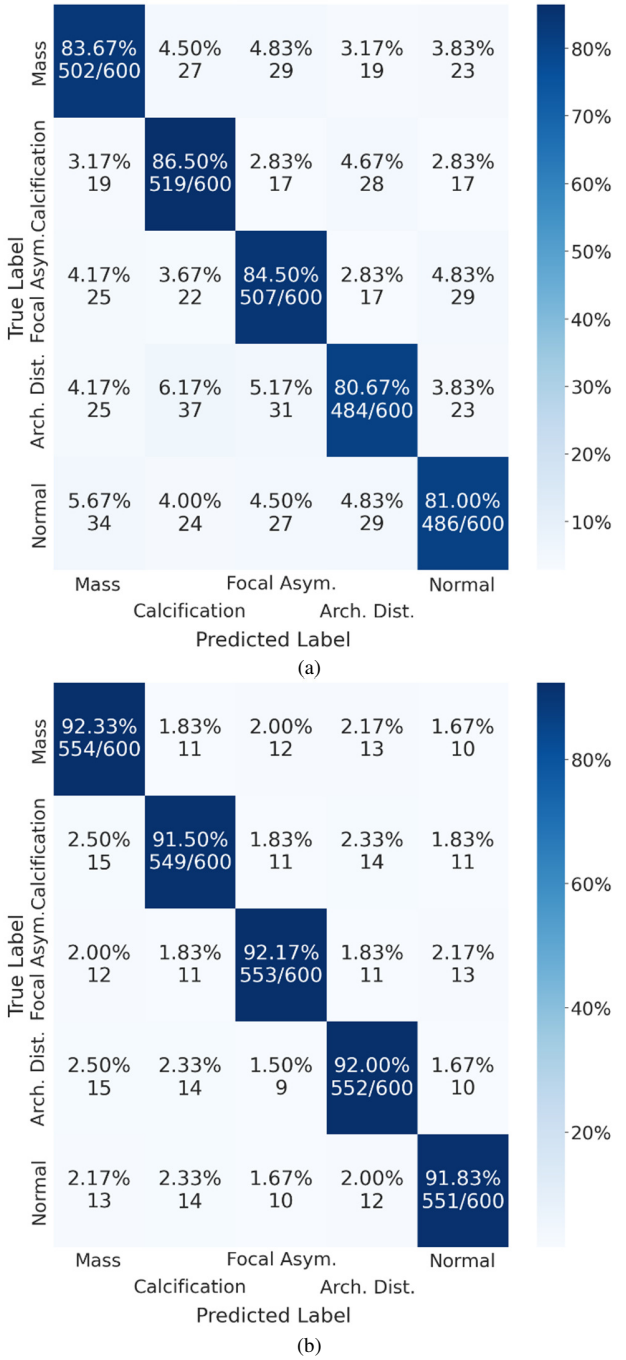


Fig. 10. Confusion matrices illustrating the classification performance for image sets with: (a) GROI, (b) CROI.

TABLE III. CLASSIFICATION PERFORMANCE BASED ON DIFFERENT IMAGE PROCESSING METHODS ACROSS THREE ITERATIONS (R1–R3)

Proposed method	Train Accuracy (%)				Validating accuracy (%)				F1 score (%)			
	R1	R2	R3	Avr	R1	R2	R3	Avr	R1	R2	R3	Avr
Original images	75.1	75.9	77.3	76.1	72.6	73.1	72.6	72.8	72.7	73.8	74.5	73.7
GROI	86.6	89.2	88.1	88.0	82.2	83.7	83.5	83.1	82.2	83.6	84.1	83.3
CROI	94.1	93.2	94.4	93.9	91.9	91.9	92.7	92.2	91.6	91.7	92.5	91.9

TABLE IV. PRECISION, RECALL AND F1 SCORE FOR EACH IMAGE CLASS WHEN USING CROI INPUTS

Type	Precision (%)	Recall (%)	F1 score (%)
Mass	90.97	92.33	91.65
Calcification	91.65	91.50	91.58
Focal asymmetry	92.94	92.17	92.55
Architectural distortion	91.69	92.00	91.85
Normal	92.61	91.83	92.22
Average	91.97	91.97	91.97

In [35], a breast ultrasound dataset was utilized with the modern DeepLabV3+ architecture to segment breast lesions. To enhance the representation of informative features, the study proposed a modified DeepLabV3+ model by incorporating the Convolutional Block Attention Module (CBAM) into both the encoder and decoder components. A comparative analysis was conducted between the original and the modified DeepLabV3+ models using performance metrics such as dice coefficient, Intersection over Union (IoU), precision, recall, and specificity. The modified model demonstrated superior performance achieving precision, recall, specificity, dice coefficient, and IoU values of 0.974, 0.933, 0.997, 0.951, and 0.933, respectively.

A previous study used the MIAS image set, which contains three types of breast lesions: fatty, dense-glandular, and fatty-glandular. The ROIs were extracted by identifying abnormal regions within the mammograms [36]. Then, the image sets with ROIs were used to train a ResNet with 50 layers. The results of this study show that the average classification accuracy was 97.81% after 70% of the training process, 98% after 80%, and reached the optimal value after 90%.

In our study, five categories of mammography image sets were segmented using the Otsu algorithm for extracting the GROIs and these image sets were converted into images with the CROI. For evaluating the efficiency of the CROI images EfficientNet-B7 was applied, achieving an average F1 score of 91.9%, as show in Table III. The results indicate that the proposed method is highly effective compared to previous methods.

However, it is important to note that the dataset used in our study differs from those used in previous works. Therefore, our performance evaluation focuses only on the proposed algorithms. Specifically, we used five types of the breast lesion images from the VinDr-Mammo dataset, whereas previous researches often used other datasets, such as MIAS or CBIS-DDSM, which include three classes: normal, benign, and malignant. The accuracy of our proposed approach may be affected by datasets with low-quality, noisy input images. Moreover, the algorithm must be evaluated using different network models to determine its application range and compatibility. These aspects will be explored in future studies.

IV. CONCLUSION

This study introduced a novel image preprocessing pipeline for enhancing breast cancer classification from mammography images, specifically addressing the imbalanced nature of the VinDr-Mammo dataset. Our approach involved two key algorithms: an optimal adaptive Otsu segmentation for

extracting the Gray Region of Interest (GROI) and an Identification and Comparison (IaC) algorithm for converting GROI into Color Region of Interest (CROI). To counter data imbalance, we extensively augmented the dataset using rotation and flipping, ensuring that each of the five lesion categories contained 3000 images, thus providing a robust foundation for deep learning.

We rigorously evaluated the effectiveness of our proposed methods by classifying breast images in three states—original, GROI, and CROI—using a fine-tuned EfficientNet-B7 model. The results demonstrated the superior performance of CROI images, achieving an average classification accuracy of 92.6%. This significantly outperformed both the original images and GROI images. Compared to a recent study using the same VinDR-Mammo dataset, which directly used grayscale images with the YOLOv5 model and achieved a peak accuracy of 81% [37], our proposed method shows better performance, affirming the efficacy of our preprocessing algorithms in highlighting critical features for breast cancer detection.

The success of our methodology, particularly with CROI, highlights its potential for broader application in classifying diverse breast cancer image sets and for integration into real-world diagnostic systems. Future work will explore the generalizability of these algorithms across different image datasets and investigate their compatibility with a wider array of advanced deep learning architectures to further enhance classification performance and adaptability for clinical applications.

ACKNOWLEDGMENT

We would like to thank the Ho Chi Minh City University of Technology and Education (HCMUTE), Vietnam.

REFERENCES

[1] B. Smolarz, A. Z. Nowak, and H. Romanowicz, "Breast Cancer—Epidemiology, Classification, Pathogenesis and Treatment (Review of Literature)," *Cancers*, vol. 14, no. 10, May 2022, Art. no. 2569, https://doi.org/10.3390/cancers14102569.

[2] M. Arnold *et al.*, "Current and future burden of breast cancer: Global statistics for 2020 and 2040," *The Breast*, vol. 66, pp. 15–23, Dec. 2022, https://doi.org/10.1016/j.breast.2022.08.010.

[3] J. D. Fackenthal and O. I. Olopade, "Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations," *Nature Reviews Cancer*, vol. 7, no. 12, pp. 937–948, Dec. 2007, https://doi.org/10.1038/nrc2054.

[4] A. Bhushan, A. Gonsalves, and J. U. Menon, "Current State of Breast Cancer Diagnosis, Treatment, and Theranostics," *Pharmaceutics*, vol. 13, no. 5, May 2021, Art. no. 723, https://doi.org/10.3390/pharmaceutics13050723.

[5] S. J. S. Gardezi, A. Elazab, B. Lei, and T. Wang, "Breast Cancer Detection and Diagnosis Using Mammographic Data: Systematic Review," *Journal of Medical Internet Research*, vol. 21, no. 7, July 2019, Art. no. e14464, https://doi.org/10.2196/14464.

[6] S. M. Shaaban, M. Nawaz, Y. Said, and M. Barr, "An Efficient Breast Cancer Segmentation System based on Deep Learning Techniques," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12415–12422, Dec. 2023, https://doi.org/10.48084/etasr.6518.

[7] A. I. Dumachi and C. Buiu, "Applications of Machine Learning in Cancer Imaging: A Review of Diagnostic Methods for Six Major Cancer Types," *Electronics*, vol. 13, no. 23, July 2024, Art. no. 4697, https://doi.org/10.3390/electronics13234697.

[8] E. Justaniah, A. Alhothali, and G. Aldabbagh, "Mammogram Segmentation Techniques: A Review," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 12, no. 5, pp. 520–529, May 2021, https://doi.org/10.14569/IJACSA.2021.0120564.

[9] T.-H. Nguyen, T.-N. Nguyen, and B.-V. Ngo, "A VGG-19 Model with Transfer Learning and Image Segmentation for Classification of Tomato Leaf Disease," *AgriEngineering*, vol. 4, no. 4, pp. 871–887, Dec. 2022, https://doi.org/10.3390/agriengineering4040056.

[10] Katherine, R. Rulaningtyas, and K. Ain, "CT scan image segmentation based on hounsfield unit values using Otsu thresholding method," *Journal of Physics: Conference Series*, vol. 1816, no. 1, Feb. 2021, Art. no. 012080, https://doi.org/10.1088/1742-6596/1816/1/012080.

[11] K. Loizidou, G. Skouroumouni, C. Nikolaou, and C. Pitris, "Automatic Breast Mass Segmentation and Classification Using Subtraction of Temporally Sequential Digital Mammograms," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 10, pp. 1–11, 2022, https://doi.org/10.1109/JTEHM.2022.3219891.

[12] S. Saifullah and R. Dreżewski, "Modified Histogram Equalization for Improved CNN Medical Image Segmentation," *Procedia Computer Science*, vol. 225, pp. 3021–3030, Jan. 2023, https://doi.org/10.1016/j.procs.2023.10.295.

[13] M. Dailla, P. Kaur, and V. Dhawan, "Adaptive Gamma Correction With Weighted Distribution And Recursively Separated And Weighted Histogram Equalization: A Comparative Study," *International Journal of Engineering Research*, vol. 3, no. 8, pp. 129–133, Aug. 2014.

[14] X. Zhang, Y. Zhu, and Z. Fan, "Region of Interest Automatic Extraction for Color Image Based on Mathematical Morphology," in *2009 Ninth IEEE International Conference on Computer and Information Technology*, Oct. 2009, vol. 1, pp. 113–117, https://doi.org/10.1109/CIT.2009.103.

[15] C. Kumah, N. Zhang, R. K. Raji, and R. Pan, "Color Measurement of Segmented Printed Fabric Patterns in Lab Color Space from RGB Digital Images," *Journal of Textile Science and Technology*, vol. 5, no. 1, pp. 1–18, Jan. 2019, https://doi.org/10.4236/jtst.2019.51001.

[16] P. Oza, P. Sharma, S. Patel, F. Adedoyin, and A. Bruno, "Image Augmentation Techniques for Mammogram Analysis," *Journal of Imaging*, vol. 8, no. 5, May 2022, Art. no. 141, https://doi.org/10.3390/jimaging8050141.

[17] H. Ulutas and V. Aslantas, "A Fast and Accurate Method for Classifying Tomato Plant Health Status Using Machine Learning and Image Processing," *Elektronika ir Elektrotechnika*, vol. 29, no. 2, pp. 54–68, Apr. 2023, https://doi.org/10.5755/j02.eie.33866.

[18] D. Al-Karawi *et al.*, "A Review of Artificial Intelligence in Breast Imaging," *Tomography*, vol. 10, no. 5, pp. 705–726, May 2024, https://doi.org/10.3390/tomography10050055.

[19] B. Alane, Y. Terchi, and S. Bouguezel, "New Face Recognition System Based on DCT Pyramid and Backpropagation Neural Network," *Elektronika ir Elektrotechnika*, vol. 30, no. 1, pp. 68–76, Feb. 2024, https://doi.org/10.5755/j02.eie.35897.

[20] M. Cantone, C. Marrocco, F. Tortorella, and A. Bria, "Convolutional Networks and Transformers for Mammography Classification: An Experimental Study," *Sensors*, vol. 23, no. 3, Feb. 2023, Art. no. 1229, https://doi.org/10.3390/s23031229.

[21] Z. Zhu, Y. Sun, and B. Honarvar Shakibaei Asli, "Early Breast Cancer Detection Using Artificial Intelligence Techniques Based on Advanced Image Processing Tools," *Electronics*, vol. 13, no. 17, Sept. 2024, Art. no. 3575, https://doi.org/10.3390/electronics13173575.

[22] A. Bekkouche, M. Merzoug, M. Hadjila, and W. Ferhi, "Towards Early Breast Cancer Detection: A Deep Learning Approach," *Engineering, Technology & Applied Science Research*, vol. 14, no. 5, pp. 17517–17523, Oct. 2024, https://doi.org/10.48084/etasr.8634.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017, https://doi.org/10.1145/3065386.

[24] S.-H. Chen, Y.-L. Wu, C.-Y. Pan, L.-Y. Lian, and Q.-C. Su, "Breast ultrasound image classification and physiological assessment based on GoogLeNet," *Journal of Radiation Research and Applied Sciences*, vol. 16, no. 3, Sept. 2023, Art. no. 100628, https://doi.org/10.1016/j.jrras.2023.100628.

[25] S. Montaha *et al.*, "BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated Using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images," *Biology*, vol. 10, no. 12, Dec. 2021, Art. no. 1347, https://doi.org/10.3390/biology10121347.

[26] L. Li, C. Pan, M. Zhang, D. Shen, G. He, and M. Meng, "Predicting malignancy in breast lesions: enhancing accuracy with fine-tuned convolutional neural network models," *BMC Medical Imaging*, vol. 24, no. 1, Nov. 2024, Art. no. 303, https://doi.org/10.1186/s12880-024-01484-1.

[27] H. Li, S. Zhuang, D. Li, J. Zhao, and Y. Ma, "Benign and malignant classification of mammogram images based on deep learning," *Biomedical Signal Processing and Control*, vol. 51, pp. 347–354, May 2019, https://doi.org/10.1016/j.bspc.2019.02.017.

[28] I. Domingues, P. H. Abreu, and J. Santos, "Bi-Rads Classification of Breast Cancer: A New Pre-Processing Pipeline for Deep Models Training," in *2018 25th IEEE International Conference on Image Processing*, Athens, Greece, 2018, pp. 1378–1382, https://doi.org/10.1109/ICIP.2018.8451510.

[29] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." arXiv, Sept. 11, 2020, https://doi.org/10.48550/arXiv.1905.11946.

[30] R. Raza *et al.*, "Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images," *Engineering Applications of Artificial Intelligence*, vol. 126, Nov. 2023, Art. no. 106902, https://doi.org/10.1016/j.engappai.2023.106902.

[31] A. Abdelrahman and S. Viriri, "EfficientNet family U-Net models for deep learning semantic segmentation of kidney tumors on CT images," *Frontiers in Computer Science*, vol. 5, Sept. 2023, Art. no. 1235622, https://doi.org/10.3389/fcomp.2023.1235622.

[32] Y. Sun, Z. Zhu, and B. Honarvar Shakibaei Asli, "Automated Classification and Segmentation and Feature Extraction from Breast Imaging Data," *Electronics*, vol. 13, no. 19, Oct. 2024, Art. no. 3814, https://doi.org/10.3390/electronics13193814.

[33] H. T. Nguyen *et al.*, "VinDr-Mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography," *Scientific Data*, vol. 10, no. 1, May 2023, Art. no. 277, https://doi.org/10.1038/s41597-023-02100-7.

[34] M. R. Islam *et al.*, "Enhancing breast cancer segmentation and classification: An Ensemble Deep Convolutional Neural Network and U-net approach on ultrasound images," *Machine Learning with Applications*, vol. 16, June 2024, Art. no. 100555, https://doi.org/10.1016/j.mlwa.2024.100555.

[35] S. Aggarwal, M. Garg, A. Kumar, and R. Kapila, "Breast lesions segmentation from ultrasound images using DeepLabV3 + model with channel and spatial attention mechanism," *Discover Sustainability*, vol. 5, no. 1, Aug. 2024, Art. no. 217, https://doi.org/10.1007/s43621-024-00424-x.

[36] A. S. Elkorany and Z. F. Elsharkawy, "Efficient breast cancer mammograms diagnosis using three deep neural networks and term variance," *Scientific Reports*, vol. 13, no. 1, Feb. 2023, Art. no. 2663, https://doi.org/10.1038/s41598-023-29875-4.

[37] S. R. Kebede *et al.*, "Dual view deep learning for enhanced breast cancer screening using mammography," *Scientific Reports*, vol. 14, no. 1, Feb. 2024, Art. no. 3839, https://doi.org/10.1038/s41598-023-50797-8.

2021 International Conference on System Science and Engineering (ICSSE)

A GLCM algorithm for optimal features of mammographic images for detection of breast cancer

Thanh-Tam Nguyen
Faculty of Electrical - Electronics Engineering HCMC University of Technology and Education, Ho Chi Minh City, Vietnam
tamnt.ncs@hcmute.edu.vn

Thanh-Hai Nguyen
Department of Industrial Electronics - Biomedical Engineering
HCMC University of Technology and Education, Ho Chi Minh City, Vietnam
nthai@hcmute.edu.vn

Ba-Viet Ngo
Department of Industrial Electronics - Biomedical Engineering
HCMC University of Technology and Education, Ho Chi Minh City, Vietnam
vietnb@hcmute.edu.vn

***Abstract*—Mammography is the golden standard for imaging and diagnosis of breast cancer in early stages. However, it is difficult for radiologists to interpret results from mammograms, particularly mammogram images have low contrast and the image quality related to technician and devices. In this paper, optimal features from a Gray Level Co-occurrence Matrix (GLCM) algorithm are applied to the mammography images for increseasing the accuracy of breast cancer detection. In particular, two datasets of the mammography images are filtered and segmented to produce Glandular Tissue Region (GTR) which contain significant features. Therefore, we just choose 4 optimal features of 10 ones using the GLCM algorithm through statistic evaluation. The results show that the optimal selected features have significant impact to produce breast cancer detection with high performance using a SVM classifier.**

Keywords—Mammograms, GTR images, Optimal features using GLCM, SVM algorithm.

I. INTRODUCTION

Breast Cancer (BC) is one of cancer diseases that can occurs in both men and women, but it is far more common in women. In Globalcan statistics in 2018, BC has the third high rate (9.2%) and is behind lung and liver cancers. In [1], there were about 164,671 deaths due to BC in Vietnam. It is obvious that BC is very dangerous. Therefore, besides knowledge and experience of physicians for early diagnosis, using image processing and artificial intelligence for support in diagnosis is very necessary. This can support physicians in the sooner diagnosis of BC.

There are different medical imaging techniques [2-4] for breast cancer. Mammography [5] is popular due to being cheaper and often using for cancer screening. In addition, if there is an X-ray image without being obvious, an accurate diagnostic result is very difficult. The difficulty of exactly diagnosing is being low intensity or many similar things in the X-ray image. Therefore, image processing techniques are applied for enhancement of the image or segmentation for extracting Region of Interest (ROI) before diagnosing and showing result. The image processing techniques, which can be employed in this paper, are enhancement, filtering, segmentation, morphology, feature extraction of breast images for detecting cancer.

To diagnosing more exactly, a breast image is necessary to be enhanced and filtering the image is often applied. There are different filters such as median, mean, Gaussian and others which are often employed for smoothing or denoising [6-8]. To smoothing and removing noise in mammograms for classifying breast cancer, a mean filter was applied [9]. In this research, authors combined different filters for eliminating noise and artifacts to produce the smoothing output images for early detection of breast cancer. Therefore, enhancement of images as filtering is necessary before performing detection or classifying of cancer diseases.

A breast image with many parts on it is difficult so that physicians can make decision related to cancer soon. Therefore, processing images to separate ROI from the image using segmentation methods are very necessary [10-13]. An optimized region growing technique for detecting breast masses in an image was proposed [14], in which the initial seed points and thresholds were applied to create optimization. In this research, from images with ROIs detected, features extracted using a GLCM method and a Feed Forward Neural Network (FFNN) for classifying benign and malignant. The result showed that the sensitivity of this proposed method is around 98% on 300 images for both training and testing. It is obvious that this is a good result with the proposed method, but time for training and testing should be considered.

GLCM is one of feature extraction methods, in which its features are relative to different values [15]. From ROI image, one method based on grey level co-occurrence matrix and optical density, called GLCM-OD, was proposed for extracting features of mammogram for detecting abnormal Regions [16]. With the GLCM-OD features from 358 mammographic cases, the sensitivity of the proposed system is about 97.3% with 4.9 false positives per image and the Az is 0.981. This is obvious that the GLCM is useful in extracting features for classifying or detecting types of cancer diseases.

From breast images, there are types of different artificial intelligence systems such as Neural networks, Support Vector Machine (SVM), Convolutional Neural Networks (CNNs) for classifying or detecting breast cancer [17-18]. Raw mammograms were pre-processed before features extraction using GLCM, in which each mammographic image was extracted at four different angular directions: $\theta=\{0°, 45°, 90°, 135°\}$, and two distances: $D=\{1,2\}$. Therefore, a two-stage support vector machine was applied for training and testing to classifying classes of normal, benign and malignant [19]. With this proposed method, the results showed that the sensitivity,

978-1-6654-4848-2/21/$31.00 ©2021 IEEE

specificity, positive predictive, and negative predictive values are high, around 92%. However, the number of mammograms in this research are not many to be able to evaluate exactly.

This paper is organized as follows. Section I is introduction about methods of mean filter, segmentation for extracting GTR, GLCM algorithm for extracting features and artificial intelligence for detection and classification of breast cancer. In Section II, the paper will present methodology of filtering images using mean filter, segmentation for searching GTR, GLCM algorithm for extracting optimal features, and SVM method for detection of breast cancer. The objective of Section III is that simulation results and discussion are expressed, in which a statistic of breast image datasets is performed to evaluate the accuracy of cancer images. Finally, conclusion will be shown in this Section IV.

II. METHODOLOGY

A. Average filter for removing noise

Mammograms will be resized to be the same size before performing next steps. Therefore, an average filter is applied to remove the unnecessary features and noise. The output image of the average filter $\mathbf{I}(x,y)$ is calculated using to the following formula:

$$\mathbf{I}(x,y)=\frac{1}{K}\sum_{i=-n/2}^{n/2}\sum_{j=-n/2}^{n/2}\mathbf{H}\left(i+\frac{n}{2},j+\frac{n}{2}\right)\mathbf{F}(x+i,y+j) \quad (1)$$

where **H** is the average filter transfer with the $n\times n$ size, n is odd. K is the sum of the coefficients of **H** and **F**, in which **F** is the input image,

B. Segmentation Using Gray Level Histogram Difference

Partitioning an image based on the histogram gray-level difference is the method developed from the amplitude threshold-based partitioning method [20]. For an 8-bit grayscale image, the gray histogram has 256 values, meaning that a vector has 256 dimensions. To compare two vectors, cosθ of the angle between the two vectors is calculated to provide the distance between the two vectors. If these two gray histograms are the same, then cosθ = 1, if it is orthogonal, then cosθ = 0. Therefore, we can classify parts of the mammogram based on the difference between pairs of the gray histograms. The difference d of the two vectors **u** and **v** is determined using the following formula:

$$d=\cos\left(\frac{\mathbf{u}.\mathbf{v}}{|\mathbf{u}||\mathbf{v}|}\right) \quad (2)$$

The magnitude of the difference between two gray histograms of adjacent sub-images provides a very good means of detecting gradual changes in the luminance density of mammograms. In addition, the edge between pectoral muscle and glandular tissue and between glandular tissue and image background may be blurred by noise. However, the comparison between gray histograms helps to accurately determine the edges between the pectoral muscle, glandular tissue, subcutaneous fat layer and image background.

For the effectiveness of using this method, we can build a sampling method. In particular, the original mammogram is divided into rows and columns for forming sub-images. These sub-images may overlap or not together. To accurately determine the edges between glandular tissues, pectoral muscles, fat layer, and image background, adjacent sub-images are used from left to right and from top to bottom. After the sampling procedure, the scalar product of two vectors is calculated between the gray histograms of the pairs of sub-images. This scalar product is used to construct a new image representing the changes in the luminance density on the original mammogram.

The next step is to determine the edges between the components in the new image by creating a weighted gray histogram of the original mammogram with the weights being the gray histogram difference values. All of these values are in the range of from 0 to 1. Creating the weighted gray histogram is to emphasize the intensities that appear in variable regions of the mammogram.

After having a new image from creating a weighted gray histogram, the calculation of the threshold is performed to segment this new image for separating the GTR in the new image. Let **A** be the input image vector converted to a row vector and **B** is the row vector of the image after using the weighted gray histogram. In addition, the interpolation method is applied to determine the approximation function of the vectors **A** and **B**. The result of the regression function is the matrix **R** with the 1x2 size containing the angle and freedom coefficients. The slope of the regression function is a constant considered as the threshold value T for segmentation using the following formula:

$$T=\frac{\mathbf{R_{12}}}{\mathbf{R_{11}}} \quad (3)$$

C. Feature Extraction

With 10 features obtained using the GLCM, we selected 4 features which are suitable for the dataset of mammograms used in the SVM classifier. The datasets for detection of breast cancer are extracted 4 features of 10 ones such as Contrast, Energy, Homogeneity and Mean, in which meaning and its formula are described as follows:

- Contrast shows the matrix with the elements distribution, in which each element is located far away from the main diagonal and its value of contrast is greater. The contrast value let us know the amount of the gray intensity diversity contained in the sample breast image and it is calculated using the following formula:

$$Con=\sum_{i}^{L-1}\sum_{j}^{L-1}(i-j)^2P(i,j) \quad (4)$$

in which $|i\text{-}j|$ is the grayscale difference between adjacent pixels, $P(i,j)$ is the element (i,j) of the normalized symmetrical GLCM, called the distribution probability of the different grayscale levels between the adjacent pixels. L is the number of gray levels in the mammographic image.

- The energy can be determined by measuring the image gray-scale in a breast image that reflects the distribution of weight uniformity and its texture. The energy value is expressed as follows:

$$Egy=\sum_{i}^{L-1}\sum_{j}^{L-1}P(i,j)^2 \quad (5)$$

- The Homogeneity value can describe the structure uniformity of the grayscales in a breast image and is calculated as follows:

$$Hom = \sum_{i}^{L-1}\sum_{j}^{L-1}\frac{1}{1+(i-j)^2}P(i,j) \quad (6)$$

➢ Mean is the average gray level of the mammographic image, where $G(x,y)$ is the image with the $m \times n$ size. In addition, the gray intensity of pixels is normalized in the range [0,1] before calculating the average of mammographic image as follows:

$$Me = \frac{\sum_{x=0}^{m-1}\sum_{y=0}^{n-1}G(x,y)}{(2^{16}-1)\times m \times n} \quad (7)$$

D. SVM Classification

For detection of breast cancer, a SVM is applied, in which 4 types of the optimal features are used to train and test the input of the SVM. This classifier system allows to be able to detect breast cancer. From features extracted using the GLCM, the linear hyperplane in the SVM algorithm is determined to divide the feature datasets for training and classifying to detect breast cancer.

For classifying breast cancer using the SVM algorithm [21], a Lagrange multiplier is necessary. In this algorithm, the Lagrange multipliers α_i, with *i = 1, 2, ..., m* correspond to the inequality constraints and their Lagrangian relating to the normal vector **w** and the feature vector **v** and it is represented as follows:

$$L(\mathbf{v},\mathbf{w},\alpha) \equiv \frac{1}{2}\|w\|^2 - \sum_{i=1}^{p}\alpha_i y_i(\mathbf{w}^T v_i + b) + \sum_{i=1}^{p}\alpha_i \quad (8)$$

In this algorithm, each training sample corresponds to v_i and the Lagrange multipliers α_i. Thus, after training, $\alpha_i \geq 0$ is the support vector located on one of the two hyperplanes.

III. Results and Discussion

A. Digital mammogram database

The image data exploited in this research was obtained from the Mini-DDSM datasets [22]. The mammograms adopted is in the form of 16-bit gray level images. The image format is in the form of Portable Network Graphics, including 9684 images of 3 types (Benign, Cancer and Normal). In particular, this paper used the part of the Mini-DDSM datasets for processing, extracting features and classifying breast cancer as described in Table I:

Table I. Representation of datasets for training and testing of 2 types of mammograms

Training and testing images	Cancer	Normal
Training	100	100
Testing	50	50

B. Extraction of Region of Interest

The original image as shown in Figure 1 after normalization is applied to an averaging filter and its result is described in Figure 2, in which the small features were removed. In addition, the method of finding the gray level difference from a histogram is applied to enhance the filtered image and its result is as shown in Figure 3. the threshold detection algorithm for segmentation of the enhanced image is based on the path approximation of the gray histogram, in which the weights are performed and Figure 4 is the result after segmentation. It is obvious that the segmentation is to remove pectoral muscle and extraneous parts for obtaining GTR of the glandular tissue.

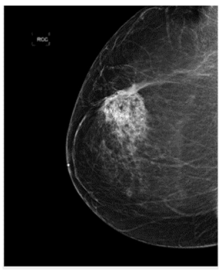


Figure 1. Original breast image

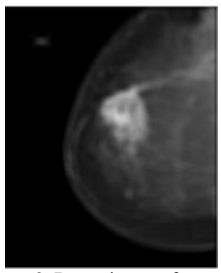


Figure 2. Breast image after using the average filter

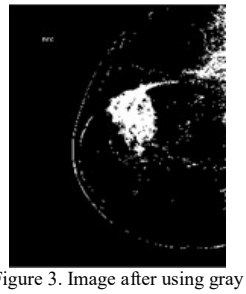


Figure 3. Image after using gray level difference

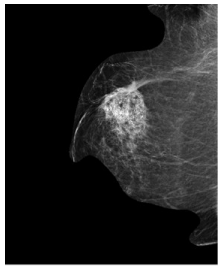
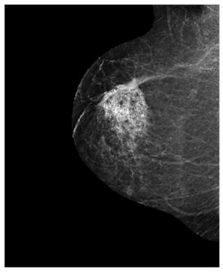


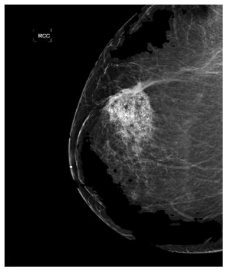
Figure 4. Image after segmenting to separate glandular tissue

Figure 5 is the segmentation result with different thresholds T, in which the thresholds are T = 9,3260 (Figure 5a), T = 29,3260 (Figure 5b) and T = 0.93260 (Figure 5c). In these results, there is a poor segmentation result, with the threshold T = 19,3260 (Figure 5d) using the Histogram Difference method.

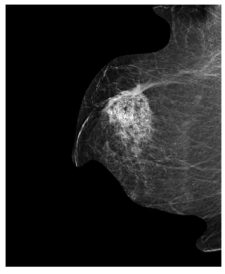
(a) T = 9.3260



(b) T = 29.3260



(c) T = 0.93260

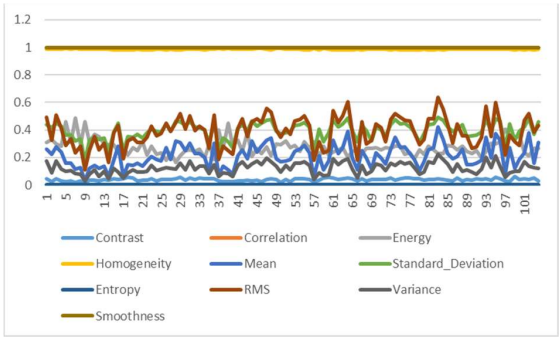


(d) T = 19.3260

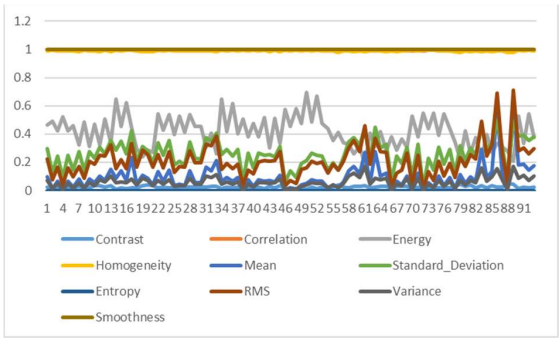
Figure 5. Segmentation of breast images with different thresholds T

C. Feature Extraction Using GLCM

Breast images containing GTR are analyzed for feature extraction using the GLCM and this allows to detect latent cancers in breasts. Therefore, the GLCM features are employed for training and classifying using a SVM method. The feature extraction procedure is based on statistics for finding the best group of features for the SVM classification.



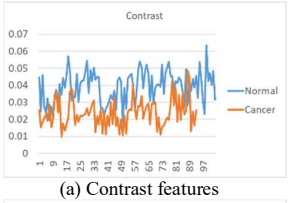
(a) 10 features of the normal images



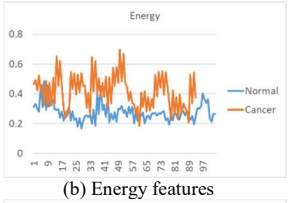
(b) 10 features of the cancer images

Figure 6. Feature statistics of breast images using GLCM

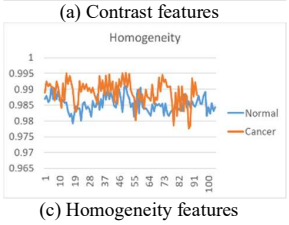
From Figure 6, it can be seen that from the statistics of the GLCM feature extraction results for 10 types of 300 breast images, 4 optimal features are found in 10 feature types, including Contrast, Energy, Homogeneity and Mean as described in Figure 7. In addition, these features show best description of the difference between lesion and non-lesion areas in breast images. Therefore, the 4 optimal features are the inputs of the SVM classifier for classifying normal and cancer cases.



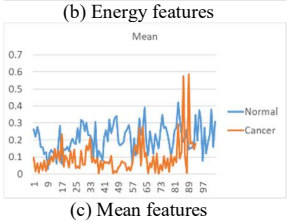
(a) Contrast features



(b) Energy features



(c) Homogeneity features



(c) Mean features

Figure 7. Optimal features using GLCM of two types of Normal and Cancer images

D. Results of SVM Classification

In the testing step after the SVM classification for the tested data, the calculation to find out the accuracy and sensitivity of the system is performed using the confusion matrix as shown in Figure 8. With the cancer image dataset of 50 images, 45 images were used for classifying and it produced the accuracy of 90%. Furthermore, with 50 normal images, the system was classified to produce the accurate result of 41 images, accounting for 82%. Therefore, it means that the average accuracy for both types was 86%. In [23], the research used 4 types of features including Contrast, Correlation, Energy, Entropy and the SVM classifier applied on the MIAS (mammographic Image Analysis Society) image set with the average accuracy of 88.2%. In particular, the accuracy with 2 types of normal and cancer is 100% and 69%, respectively. Although the average accuracy is 2.2% lower than that of the research in [23], the proposed method gives 21% higher ability to classify cancer images. Moreover, the study in [24] used the GLCM feature extraction method and the SVM classifier with the accuracy of just 74.5%.

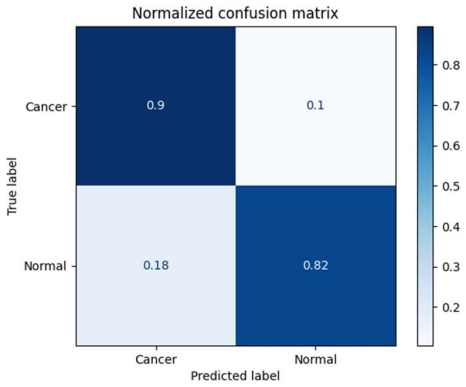


Figure 8. Confusion matrix for evaluation of 50 testing images each class

IV. CONCLUSION

In this paper, we studied the problem of classifying breast images for detecting cancer and evaluating its accuracy. The breast image datasets were re-sized and filtered to enhance them before segmentation to produce GTR areas. In this research, the GLCM was applied to extract 4 optimal features which were performed based on the statistic of 10 features for classifying. The SVM classifier was employed to classify 2 types of breast images (normal and cancer) and produced the significant accuracy evaluated using the confusion matrix. This result shows that the proposed method is effective with the datasets of just 300 breast images. In addition, the processing breast images with proposed method can be developed to larger datasets and other classifiers.

ACKNOWLEDGMENT

We would like to thank Ho Chi Minh City University of Technology and Education (HCMUTE), students and colleagues for supports on this paper.

REFERENCES

[1] World Health Organization, “Vietnam Source: Globocan 2018,” 5-2019

[2] Chen, Biao, and Ruola Ning, "Cone-beam volume CT breast imaging: a Feasibility study," Medical Physics, vol. 29, no. 5, pp. 755-770, 2002

[3] Jalalian, A., Mashohor, S. B., Mahmud, H. R., Saripan, M. I. B., Ramli, A. R. B., & Karasfi, B., "Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review," Clinical imaging, vol.37, no. 3, pp. 420-426, 2013

[4] Mann, R. M., Kuhl, C. K., Kinkel, K., & Boetes, C.,"Breast MRI: guidelines from the European society of breast imaging," European Radiology, vol.18, no. 7, pp. 1307-1318, 2008

[5] Olsen, Ole, and Peter C. Gøtzsche., "Cochrane review on screening for breast cancer with mammography," The Lancet, vol. 358, no. 9290, pp. 1340-1342, 2001

[6] Ardra Mariya Joseph, M. Grace John, Anto Sahaya Dhas, "Mammogram image denoising filters: A comparative study," the Conference on Emerging Devices and Smart Systems (ICEDSS), 2017

[7] R. Ramani, Suthanthira Vanitha, S. Valarmathy, "The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images," the International Journal of Image Graphics and Signal Processing, vol. 5, no. 5, pp. 47-54, 2013.

[8] Bijo John, Saravanan Nallathambi, "Study and Analysis of Filters," Advances in Computational Sciences and Technology, vol. 10, no. 3, pp. 331-341, 2017.

[9] Shailaja Singh1, Anamika Yadav2, and Bikesh Kumar Singh., "Performance Analysis of Mammographic Image Enhancement Techniques for Early Detection of Breast Cancer," International Conference on Parallel Distributed Computing Technologies and Applications, Advances in Parallel Distributed Computing, pp 439-448, 2011.

[10] Zhang, Jun & Hu, Jinglu, "Image segmentation based on 2D Otsu method with histogram analysis," 2008 International Conference on Computer Science and Software Engineering, vol. 6, pp. 105–108, 2008

[11] Mustra, Mario, Mislav Grgic, and Rangaraj M. Rangayyan, "Review of recent advances in the segmentation of the breast boundary and the pectoral muscle in mammograms," Medical & biological engineering & computing, vol. 54 no. 7, pp.1003-1024, 2016

[12] Shi, P., Zhong, J., Rampun, A., & Wang, H., "A hierarchical pipeline for breast boundary segmentation and calcification detection in mammograms," Computers in biology and medicine, vol. 96, pp. 178-188, 2018

[13] Kashyap, K. L., Bajpai, M. K., Khanna, P., Giakos, G., "Mesh-free based variational level set evolution for breast region segmentation and abnormality detection using mammograms," International journal for numerical methods in biomedical engineering, vol. 34, no. 1, 2018

[14] S.PunithaaA.AmuthanbK. SureshJosepha, "Benign and malignant breast cancer segmentation using optimized region growing technique," Future Computing and Informatics Journal, vol. 3, issue 2, pp. 348-358, 2018.

[15] A Mohd. Khuzi, BEng, R Besar, WMD Wan Zaki, MEngSc, NN Ahmad, "Identification of masses in digital mammogram using gray level co-occurrence matrices," Biomedical Imaging and Intervention Journal, vol. 5, no. 3, pp. 1-13, 2009

[16] Zih-Siou Chen, Wei-Ting Tsai, Chin-Peng Lin, and Li-li Cheng, "A Mass Detection System in Mammograms Using Grey Level Co-occurrence Matrix and Optical Density Features," Advances in Intelligent Systems & Applications, vol. 2, pp. 369–376, 2013.

[17] Syam Julio A. Sarosa, Fitri Utaminingrum, Fitra A. Bachtiar, "Mammogram Breast Cancer Classification Using Gray-Level Co-Occurrence Matrix and Support Vector Machine," 2018 International Conference on Sustainable Information Engineering and Technology (SIET), 2019

[18] Youssef Ben Youssef, Elhassane Abdelmounim, Abdelaziz Belaguid, "Mammogram Classification Using Support Vector Machine," Handbook of Research on Advanced Trends in Microwave and Communication Engineering, chapter 9, 2017.

[19] Aderonke Anthonia Kayode, Noah Oluwatobi Akande, Adekanmi Adeyinka Adegun, Marion Olubunmi Adebiyi, "An automated mammogram classification system using modified support vector machine," Med Devices, vol. 12, pp. 275-284, 2019.

[20] [24] Nguyen Thai Ha, Nguyen Duc Thuan, Pham Manh Hung, Dao Trang Linh, Tran Thanh Minh, "Segmenting Mammogram By Histogram Difference Method," Journal of Science and Technology Technical Universities, vol. 74, pp. 67-71, 2013.

[21] Thanh-Hai N, Hoai-An T, "PCA-SVM Algorithm for Classification of Skeletal Data-Based EigenPostures," American Journal of Biomedical Engineering, vol. 6, pp. 147-158, 2016.

[22] C.D. Lekamlage, F. Afzal, E. Westerberg and A. Cheddad, "Mini-DDSM: Mammography-based Automatic Age Estimation," International Conference on Digital Medicine and Image Processing (DMIP 2020), 2020.

[23] E. Palantei, Asma Amaliah, Indrabayu Amirullah, "Breast Cancer Detection in Mammogram Images Exploiting GLCM, GA Features and SVM Algorithms," Journal of Telecommunication, Electronic and Computer Engineering, vol. 9, pp. 113-117, 2017.

[24] Htay T T and Maung S S, "Early Stage Breast Cancer Detection System using GLCM feature extraction and K-Nearest Neighbor (k-NN) on Mammography image," International Symposium on Communications and Information Technologies (ISCIT), pp. 171-175, 2018

Breast Image Segmentation for evaluation of Cancer Disease

Thanh-Tam Nguyen
Faculty of Biomedical Engineering
International University, VNU,
Ho Chi Minh City, Vietnam
nttam@hcmiu.edu.vn

Thanh-Hai Nguyen
Faculty of Electrical and Electronics Engineering
Ho Chi Minh City University of Technology and Education,VietNam
nthai@hcmute.edu.vn

Ba-Viet Ngô
Faculty of Electrical and Electronics Engineering
Ho Chi Minh City University of Technology and Education, Vietnam
vietnb@hcmute.edu.vn

Duc-Dung Vo
Faculty of Electrical and Electronics Engineering
Ho Chi Minh City University of Technology and Education,VietNam
dungvd@hcmute.edu.vn

***Abstract*—Breast cancer is one of dangerous diseases and difficult to cure. It is observed that early detection of malignancy can help in the diagnosis of the disease and patient can be saved. For the detection of breast cancer, breast images will be enhanced using a Fuzzy logic and possibility distribution algorithm and then segmented to produce images with region of interest, in which just cancer shape appears in the image for detecting and estimating disease status. This paper proposes a statistic method based on the gray level of pixels in the image through histograms of two breast image sets to classify two cases of cancer and normal one. Simulation results on breast image sets will show that the proposed method is effective and it can be developed for detection of benign and malignant tumors in artificial intelligent systems.**

Keywords— Breast image sets; Image enhancement; Image segmentation; Histogram for evaluation; Statistics of contrast

I. INTRODUCTION

Introduction.

Cancer is one of the most dangerous and very difficult diseases to treat for one patient. In types of cancer, Breast Cancer (BC) often appears in women, the most women are 40 years old. In current, there are about 164,671 breast cancer diseases in Vietnam and about 114,871 people were died [1]. In addition, BC has the third high rate (9.2%) and is behind liver and lung cancers. In particular, BC in women accounted about 15,229 of the 73,849 cancer cases according to Globalcan statistics in 2018. Therefore, BC is possibly diagnosed early, the chance of complete cure is very high. Thus, using methods of image processing for analysing and diagnosing diseases early are very important.

Diagnostic imaging doctors often use various imaging methods including Computerized Tomography (CT) [2], ultrasound [3], mammography [4] and Magnetic Resonance Imaging (MRI) [5] for screening and diagnosis to detect cancer disease early. In these methods, ultrasound and x-ray images mainly used to detect and diagnose breast cancer. The blurred image edges and the low contrast of the ultrasound image are one challenge in automatic image segmentation. During capturing x-ray breast image, high-resolution images with low-energy X-rays allow to detect abnormalities or tumors obscured or overlapped by surrounding breast tissue [6-7]. To extract anomalies or areas of interest from x-ray breast images, image segmentation can be applied. In practice, there are various segmentation and in diagnostic techniques, pre-processing needs to be applied to remove labels, tags, patient names or other unwanted information. In addition, these techniques increase image contrast and eliminate noise for making the segmented images more accurate and reliable.

After image preprocessing, image segmentation algorithms play an important role in determining whether tumor in image is malignance or benign. To detect malignant tumor, features such as intensity, shape, size, texture, gray scale histogram for describing the tumor can be calculated. These segmentation algorithms can be classified into groups including regional approach (grouping pixels into homogeneous regions during large computation based on high resolution) [8]; based on contours (meaning based on discontinuity of color, gray level or texture of image edges detected respectively), based on cluster (pixel clusters with the same property) [9]; threshold method (foreground segmentation from background through information from gray level histogram); methods based on energy function [10].

Region-based algorithm applied on mini-MIAS database for chest muscle segmentation was the 98% accuracy [11-12] and in the case of EPIC data sets, this algorithm had the 91.5% accuracy. The accuracy of this algorithm in removing noise and extraneous components from the mini-MIAS dataset was very good, about 98.8% [13]. However, this technique is mostly used due to its high resolution, so it takes more time during segmentation as well as requiring the selection of same points is difficult. Another technique is based on analyzing the energy of components in the image to eliminate unnecessary components for enhancing the image and removing chest muscle from image, with the 90.37% accuracy [14]. This technique is very flexible and just requires little calculation.

Clustering techniques were proposed to apply in methods of SVM, fuzzy c-means and decision trees [15-21], in which they could work well for overlapping data, giving high accuracy in clustering and detecting breast tumors from images having sensitive problems to initial clusters and peripheral values. In addition, clustering and texture filters can effectively detect calcification even with small noise image or large tumors inside the breast. Moreover, the c-mean fuzzy clustering technique has large tolerance for contrast, fuzzy boundaries, noise and this can produce high precision in image segmentation. Therefore, this method is better for segmenting breast tumor lesions.

This paper is organized as follows. Section I is introduction about methods of enhancement, segmentation, binary conversion related to classification of breast images. In Section II, the paper will present methods of enhancement, segmentation for searching ROI, binary conversion and gray level statistics for searching features of breast images. The objective of Section III is that simulation results and discussion are expressed, in which a statistic of breast image sets is produced to evaluate cancer images. Finally, conclusion will be shown in this Section IV.

II. Material And Method

A. Algorithm for Image Enhancememt

Breast cancer disease is very dangerous and very difficult to detect it in the first stage. This difficulty can be due to two reasons: patient is not regularly screened to detect cancer and breast image sometimes does obvious for diagnosis. Therefore, processing breast images to detect breast cancer soon is very necessary. In this paper, the source of datasets was collected from Mammographic Image Analysis Society (MIAS). Original images are often hard to accurately diagnose breast image status with cancer or normal. Therefore, the images can be enhanced before segmentation for detecting Region of Interest (ROI) and then statistics for evaluation of disease status as shown in Fig. 1.

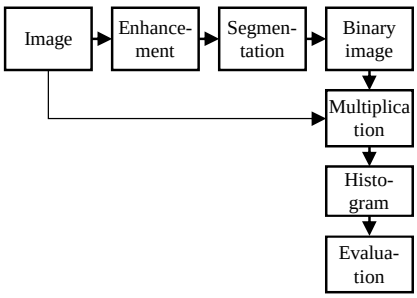


Fig. 1. Block diagram of cancer evaluation process

After collecting the datasets from MIAS, the input images are enhanced using a Fuzzy logic and possibility distribution algorithm [21]. In the research, the minimum, maximum and mean values of pixels of one gray level image are calculated and then two thresholds of Th_1, Th_2 are calculated as follows:

$$Th_1 = (mean + min)/2 \quad (1)$$

$$Th_2 = (mean + max)/2 \quad (2)$$

From the two threshold values, the gray level image is divided into four groups for determining four corresponding thresholds as follows:

$$min \le P_0 < Th_1$$

$$Th_1 \le P_1 < mean$$

$$mean \le P_2 < Th_2$$

$$Th_2 \le P_3 \le max$$

in which,

- P_O, P_1, P_2, P_3 are the new gray level thresholds
- *min* denotes the smallest gray level value in the image
- *max* is the largest gray level value
- *mean* describes the gray level average value

The gray level of pixels in each group is adjusted and then all pixels in the image are calculated to correspond to gray levels using the following equations:

$$P_{N0} = 2 \times \left(\frac{(P_0 - min)}{(mean - min)}\right)^2 \quad (3)$$

$$P_{N1} = 1 - 2 \times \left(\frac{(P_1 - mean)}{(mean - min)}\right)^2 \quad (4)$$

$$P_{N2} = 1 - 2 \times \left(\frac{(P_2 - mean)}{(max - min)}\right)^2 \quad (5)$$

$$P_{N3} = 2 \times \left(\frac{(P_3 - mean)}{(max - mean)}\right)^2 \quad (6)$$

where P_{N0}, P_{N1}, P_{N2}, P_{N3} are four new thresholds calculated based on previous thresholds and the values of *min*, *max* and *mean* in the breast image.

From Eq. (3) to Eq. (6), all breast images are enhanced and contrast levels are adjusted. It is obvious that the breast images after enhancement for image segmentation are better to search ROI areas.

B. Otsu Image Segmentation

After enhancement of images, an Otsu segmentation algorithm is applied to determine thresholds of gray level for convert gray level images to binary images [22]. In particular, the Otsu segmentation is to determine gray level, where a gray level image is calculated to divide pixels into two groups: background pixels, and object pixels. Moreover, the threshold is calculated to minimize the intra-class variance of two classes (background and object). In order to search this threshold, the algorithm is described as follows:

1. Calculate histogram and probabilities of each class in image
2. Set up initial probabilities
3. Update probabilities and calculate mean values of two classes. Step through all possible
4. Desired threshold corresponds to the maximum value of

The Otsu threshold is applied to convert the gray level image into a binary image, in which the ROI area is 1s pixels and the background is 0s pixels. For representation of ROI histogram, the binary image is multiplied to the original image to produce the gray level ROI. Therefore, the probability density of pixels in the image with the gray ROI will produce a gray level histogram, which allows us calculate the different gray level area between two types of breast image (cancer and normal). From these different gray level areas of the cancer and normal images, we can calculate to determine breast cancer image.

C. Determination of Breast Image

In this paper, statistics of contrast and corresponding pixel amount in two breast image sets are performed based on gray level histograms of two types of breast image (cancer and normal). These statistic represents two types of gray level corresponding to two groups of cancer and normal breast images which can be considered as features of these types. Therefore, we can determine an ability of breast cancer disease based these features. The contrast C of gray level ROI and the corresponding pixels in a breast image sets are determined by using the following formulas:

$$C = \frac{1}{NM}\sum_{i=1}^{M}\sum_{j=1}^{N}[[i-j]^2 p(i,j)] \quad (7)$$

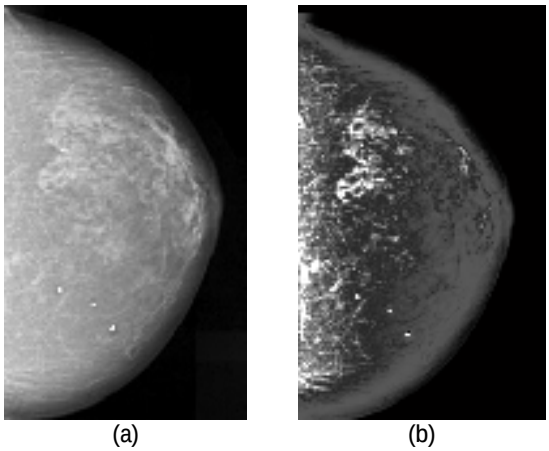
in which N, M are the dimensions of the Gray-Level Co-Occurrence Matrix (GLCM) of the image, and $p(i,j)$ is the frequency related to the gray levels i, j of two adjacent pixels.

III. Results and Discussion

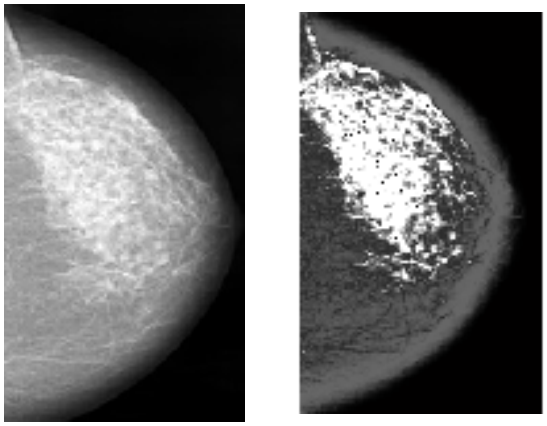
In this research, simulation results are worked out from original breast image sets using algorithms of enhancement, segmentation, binary image conversion and statistics based on gray level histograms, in which the breast image sets consist of ten normal images and ten cancer ones.

A. Results of Image Enhancement

For enhancement of a breast image, ***min***, ***max*** and ***mean*** values in each image are calculated for determining thresholds of P_{N0}, P_{N1}, P_{N2}, P_{N3}. Fig.2 shows original and enhanced breast images. It is obvious that the breast images after enhancing show gray level areas which we can identify cancer status through next processing.



(a) (b)

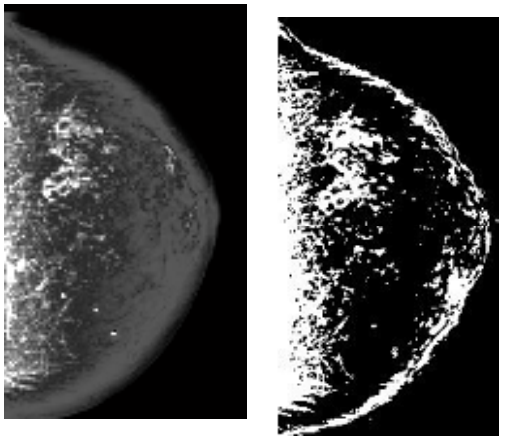


(c) (d)

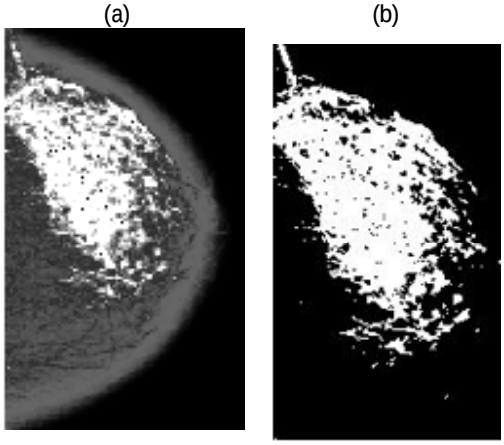
Fig. 2. Representation of original and enhanced images
(a) original image of normal case
(b) enhanced image of normal case
(c) original image of cancer case
(d) enhanced image of cancer case

B. Results of Image Segmentation

From enhanced breast images, the Otsu segmentation method was applied to produce threshold which allow to separate ROI for evaluation of image status. Breast images after segmentation were converted into binary images as shown in Fig. 2. With these binary images, basically we can see the difference between white areas (0 gray level) and back ones (1 gray level) in two cases of cancer and normal.



(a) (b)



(c) (d)

Fig. 3. Representation of enhanced and segmented images
(a) enhanced image of normal case
(b) segmented image of normal case
(c) enhanced image of cancer case
(d) segmented image of cancer case

C. Breast Image Evaluation

For evaluation of cancer status, binary images were multiplied with original breast images to produce images with gray level ROI which allow to identify between cancer and normal images. Therefore, probabilities of gray level pixels were calculated to create a gray level histogram for breast image evaluation.

From breast images with the ROI, all breast images were calculated to create gray level histograms. Therefore, the gray level densities between two types breast images (cancer and normal) were different as shown in Fig. 4 and Fig.5. In particular, the pixel density of cancer images has the sudden change in the gray level range of around 0.18-0.20, while that of normal images slowly changes in the same range.

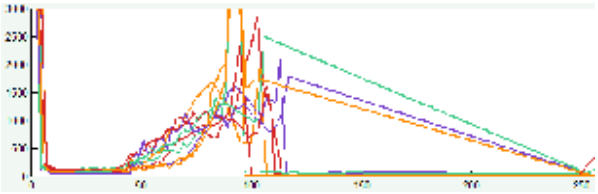


Fig. 4. Gray level histograms of normal case

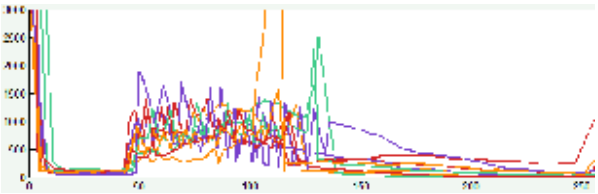


Fig. 5. Gray level histogram of cancer case

Fig. 4 and Fig. 5 show that in cancer cases, pixels corresponding to the gray level range of around 50 to 130 in the histograms change suddenly from low to high, while those corresponding to the same gray level range in normal cases increase slowly. Therefore, we can basically evaluate the difference between two types of breast images. However, contrast values and corresponding pixel amount provided more information for evaluation of breast disease status.

TABLE I. FEATURER EXTRACTION OF CANCER AND NORMAL CASES

Case	*Cancer*		*Normal*	
	Contrast	***ROI pixels***	***Contrast***	***ROI pixels***
1	0.1927	558	0.0758	319
2	0.1478	417	0.0711	331
3	0.1582	484	0.0702	207
4	0.1419	402	0.0269	343
5	0.1501	601	0.0458	391
6	0.0817	513	0.0181	212
7	0.2799	527	0.0584	388
8	0.2483	560	0.0855	326.
9	0.1185	442	0.0746	98
10	0.1455	420	0.0295	77

Table 1 shows the contrast values corresponding to the average pixel amount of ROIs in cancer and normal images in the gray level range of 0.2-0.3 in the histograms. It is obvious that the contrast values in the normal images are from 0.0181 to 0.0855, while those in the cancer images are from 0.0817 to 0.2799. In addition, ROI pixels corresponding to image contrast of ten ROI images for each type were different. Particular, ROI pixels in cancer images were the range of 402 to 601, while the range of ROI pixels in normal images was 77 to 391.

With the experiment of 20 breast images including ten image for each type, one can evaluate the ability of disease status based on image after processing. In particular, with this statistic table, we can evaluate breast image status and this can be one of information for doctor in disease diagnosis.

Mammography images were segmented using fuzzy c-means clustering and then ROIs in these image were calculated for determining normal and abnormal regions [21]. Result of the classification accuracy in this research was around 92%. In our research, image processing techniques were applied, histograms of ROIs in mammography images were represented and then statistics of pixel densities on both normal and abnormal images were performed to be the basic for evaluation. Contrast values and corresponding pixel amount were determined for evaluation of disease status.

IV. CONCLUSION

In this paper, the set of 20 breast images including 10 image for each type were processed to produce features for evaluation of disease status. Therefore, this paper presented the proposed method consisting of the image enhancement using the Fuzzy logic and possibility distribution algorithm, Otsu segmentation and statistics on ROI image based on histograms for evaluation of breast cancer. In particular, from cancer and normal images, features based on contrast values and the corresponding pixel amount were obtained as shown in Table 1, in which the ranges from the minimum and maximum values between two types of normal and cancer images were really obvious. Simulation results showed the effectiveness of the proposed method and also it is very significant for development of breast image recognition using artificial intelligence with higher performance.

ACKNOWLEDGMENT

The authors would like to acknowledge the support by the HCMC University of Technology and Education, Vietnam.

REFERENCES

[1] “Vietnam Source: Globocan 2018”, World Health Organization, 5-2019.

[2] Chen, Biao, and Ruola Ning "Cone-beam volume CT breast imaging: a Feasibility study." Medical Physics 29.5 (2002): 755-770.

[3] Jalalian, A., Mashohor, S. B., Mahmud, H. R., Saripan, M. I. B., Ramli, A. R. B., & Karasfi, B. "Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review," Clinical imaging 37.3 (2013): 420-426.

[4] Olsen, Ole, and Peter C. Gøtzsche. "Cochrane review on screening for breast cancer with mammography" The Lancet 358.9290 (2001): 1340-1342.

[5] Mann, R. M., Kuhl, C. K., Kinkel, K., & Boetes, C."Breast MRI: guidelines from the European society of breast imaging" European Radiology 18.7 (2008): 1307-1318.

[6] Johns, Paul C., and Martin J. Yaffe. "X-ray characterization of normal and neoplastic breast tissues" Physics in Medicine & Biology 32.6 (1987): 675.

[7] Mustra, Mario, Mislav Grgic, and Rangaraj M. Rangayyan. "Review of recent advances in the segmentation of the breast boundary and the pectoral muscle in mammograms" Medical & biological engineering & computing 54.7 (2016): 1003-1024.

[8] Szeliski, Richard "Computer vision: algorithms and application" Springer Science & Business Media, 2010.

[9] Muthukrishnan, R., and Miyilsamy Radha "Edge detection techniques for image segmentation," International Journal of Computer Science & Information Technology 3.6 (2011): 259.

[10] Kas, Michael, Andrew Witkin, and D. Terzopoulous. "Snakes-active contours models." International Journal of Computer Vision 1.4 (1987): 321-331.

[11] Raba, D., Oliver, A., Martí, J., Peracaula, M., & Espunya, J. "Breast segmentation with pectoral muscle suppression on digital mammograms," Iberian Conference on Pattern Recognition and Image Analysis. Springer, Berlin, Heidelberg, 2005.

[12] Chen, Zhili, and Reyer Zwiggelaar. "A combined method for automatic identification of the breast boundary in mammograms." Biomedical Engineering and Informatics (BMEI), 2012 5th International Conference on. IEEE, 2012.

[13] Zhang, Zhiyong, Joan Lu, and Yau Jim Yip. "Automatic segmentation for breast skin-line." the IEEE 10th International Conference on Computer and Information Technology, 2010.

[14] Lu, X., Dong, M., Ma, Y., & Wang, K."Automatic Mass Segmentation Method in mammograms based on improved VFC Snake model" Emerging Trends in Image Processing, Computer Vision and Pattern Recognition. 2015. 201-217.

[15] Anitha, J., and J. Dinesh Peter. "Mass segmentation in mammograms using a kernel-based fuzzy level set method." International Journal of Biomedical Engineering and Technology 19.2 (2015): 133-153.

[16] Touil, Asma, and Karim Kalti. "Iterative fuzzy segmentation for an accurate delimitation of the breast region." Computer methods and programs in biomedicine 132 (2016): 137-147

[17] Feng, Y., Dong, F., Xia, X., Hu, C. H., Fan, Q., Hu, Y., ... & Mutic, S."An adaptive fuzzy C-means method utilizing neighboring information for breast tumor segmentation in ultrasound images." Medical Physics 44.7 (2017): 3752-3760.

[18] Shi, P., Zhong, J., Rampun, A., & Wang, H. "A hierarchical pipeline for breast boundary segmentation and calcification detection in mammograms." Computers in biology and medicine 96 (2018): 178-188.

[19] Valdés-Santiago, D., Quintana-Martínez, R., León-Mecías, Á., & Díaz-Romañach, M. L. B."Mammographic Mass Segmentation Using Fuzzy C–means and Decision Trees." International Conference on Articulated Motion and Deformable Objects. Springer, 2018.

[20] Kashyap, K. L., Bajpai, M. K., Khanna, P., Giakos, G. "Mesh-free based variational level set evolution for breast region segmentation and abnormality detection using mammograms." International journal for numerical methods in biomedical engineering 34.1 2018.

[21] Chowdhary, Chiranji Lal, and D. P. Acharjya. "Segmentation of Mammograms Using a Novel Intuitionistic Possibilistic Fuzzy C-Mean Clustering Algorithm." Nature Inspired Computing. Springer, Singapore, 2018. 75-82.

[22] Zhang, Jun & Hu, Jinglu (2008). "Image segmentation based on 2D Otsu method with histogram analysis". Computer Science and Software Engineering, 2008 International Conference on. **6**: 105–108

JTE HCMUTE
ISSN: 2615-9740

JOURNAL OF TECHNICAL EDUCATION SCIENCE
Ho Chi Minh City University of Technology and Education
Website: https://jte.edu.vn
Email: jte@hcmute.edu.vn

Design of a Telemedicine System for Classification of Breast Cancer Images

Thanh-Tam Nguyen[1], Thanh-Hai Nguyen[1*], Tin-Trung Nguyen[2]
[1]Ho Chi Minh City University of Technology and Education, Vietnam
[2]Ho Chi Minh City Oncology Hospital, Vietnam

[*]Corresponding author. *Email: nthai@hcmute.edu.vn*

ARTICLE INFO

Received: 19/07/2025
Revised: 30/07/2025
Accepted: 21/08/2025
Published:

KEYWORDS

Telemedicine system;
Breast cancer classification;
EfficientNet-B7 model;
DICOM standard;
Protocols for information exchange.

ABSTRACT

Breast cancer is one of complex breast lesions. Therefore, accurate diagnosis to determine whether there is cancer disease or not, to determine which stage is a challenge for most doctors. This article proposes a telemedicine system for diagnosing breast cancer disease using EfficientNet-B7 in AI model, in which three image sets of Benign, Malignant and Normal are used. The main points are that, this telemedicine system is designed and calculated suitably so that a DICOM image can be transmitted from the image collected place to a server for classification and diagnosis, in which protocols and storage parts in this system are carefully selected and tested for its efficiency. Furthermore, layers and coefficients of the EfficientNet-B7 model are calculated and selected to increase the classification performance. Thus, the overall system results produced an accuracy of about 89.58%, which is a significant result for a complex and challenging system. Thus, the system can be improved in the future by enhancing the image sets, updating the deep learning network appropriately, and configuring a powerful enough server system.

Doi: https://doi.org/10.54644/jte.2025.1969

1. Introduction

Breast cancer is one of the causes leading death in women worldwide, in which an incidence is increasing, especially among young people in Vietnam. According to recent statistics [1], each year in Vietnam, approximately 22,000 people are diagnosed with the breast cancer and there are more than 9,000 deaths, accounting for 25.8% of all cancer cases in women. Moreover, according to data of Bach Mai Hospital in Vietnam, nearly 70% of patients are diagnosed at late stages (III-IV), which significantly reduces the chance of successful treatment. Therefore, the main reason may be the lack of a large-scale screening system and limited highly specialized resources. Furthermore, Bach Mai Hospital receives more than 5,000 patients per day, which puts great pressure on the team of diagnostic imaging doctors [2].

In recent years, Artificial Intelligence (AI) has been applied to remote diagnostic systems which is considered a breakthrough in the medical field [3]. In particular, recent researches from Sweden, Germany, South Korea and some other countries have demonstrated that AI not only improves the accuracy of early breast cancer detection by up to 20%, but also reduces the workload for doctors by 44% [4]. Several large Vietnamese hospitals, including Bach Mai Hospital and the Hospital of Thai Binh Medical University, have successfully implemented AI solutions like Cadai-BTM for interpreting ultrasound images. This not only help the automatically diagnostic process, also support doctors in better diagnosing X-ray or ultrasound breast images [2], [5].

Traditional diagnostic methods are mainly based on mammography or ultrasound imaging [6], and this requires doctors with high skills to be able to detect small lesions such as microcalcifications or tumors less than 1 cm in diameter [7]. However, subjectivity of doctors in image reading and overloaded working can sometimes lead to errors in diagnosing breast cancer images. In particular, a research from Lund University in Sweden indicates that about 15% to 30% of breast cancer cases have been missed in routine screening [8]. This is obvious that it is very unfortunate and dangerous for many patients.

With increasing workload and pressure to doctors at hospitals, this can lead to inaccurate diagnosis results. Therefore, a remote diagnostic system integrated with AI is a solution for improving accuracy and also reduce the workload for doctors. It means that the integration of AI into a telemedicine system is reshaping the approach to breast cancer globally. Recent researches have shown that AI not only improves diagnostic accuracy, but also provides diagnostic results of up to 94.5% compared to 88% for traditional diagnosis [9]. Furthermore, the use of the AI-based telemedicine system reduces the number of visits by approximately 40% compared to in-person visits. In Vietnam, where the rate of late detection of breast cancer is up to 70% [1], this telemedicine system promises to narrow the gap in access to quality healthcare services.

In recent years, many deep learning models for image analysis have been developed, particularly YOLO, ResNet-50 and other models have been applied for different lesion recognition. One research is that YOLOv6 combined with Federated Learning (FedL) using BreakHis and BUSI datasets achieved 98% accuracy in benign and malignant tissue classification [10]. In particular, this federated learning for model training on multiple servers without sharing patient data due to information security. Meanwhile, MammoScreen with an FDA-cleared AI solution has been employed a top-down approach for evaluating the entire mammogram and a bottom-up approach for analyzing each suspicious region, resulting in a 44% reduction in physician workload [4].

Recent healthcare systems have been developed for assisting doctors in medical diagnosis and treatment [11], [12] and AI is one of the technologies developed in some of these systems. Google Health is developing the Med-PaLM 2 system, which is a medical-specific Large Language Model (LLM) for analyzing results of test, images, and then recommending treatment regimens. In India, a trial of the IBM Watson system achieved 90% agreement with the decision of a panel of experts in the treatment of early-stage breast cancer [13]. These systems use reinforcement learning to continuously update from new clinical data for optimizing personalized treatment recommendations. In particular, Google Health's AI-powered system reduced 5.7% of false positives and 9.4% of false negatives compared to traditional methods. Another study is that with the analysis of 44,755 ultrasound images and AUROC 0.976, the system detected lesions with very small sizes from 0.5mm to 2mm that are easily missed by the human eye [14].

The EfficientNet model has been developed through several stages and the EfficientNet-B7 model achieves higher accuracy than some other deep learning models. Furthermore, EfficientNet-B7 not only produces good accuracy with image data, but also has a network size 8.4 times smaller and 6.1 times faster than some other deep learning networks [15]. In recent studies, the EfficientNet model has been combined with other algorithms for application in many different classification fields. In particular, Lung-EfficientNet was proposed for classifying lung cancer based on a set of CT scan images and produced results with an accuracy of 99.10% [16]. Another research is that EfficientNet family U-Net models have been applied for segmenting renal tumors on CT-scan images and achieved high accuracy results using tumor segmentation [17], [18]. In the telemedicine breast cancer imaging diagnosis system, we used EfficientNet-B7 for image classification for evaluating the remote classification results.

This article is organized as follows: Section 2 briefly describes the proposed methods and materials, including building hardware architecture of the telemedicine system, image processing, AI model using EfficientNet-B7, protocols for interfacing input-output blocks. In Section 3, the results of methods for classifying using this system and result evaluations. The final section is the conclusion points.

2. Materials and Methods

This article proposes a telemedicine diagnosis system using an EfficientNet-B7 model to support doctors in classifying breast cancer images. It means that the system needs to ensure the transmission and processing of standard medical images over the internet to the server to produce breast cancer image classification results. Moreover, with this telemedicine system, doctors in many remote locations can join to perform diagnosis and receive this classification result for urgent case of dangerous and complicated diseases.

Furthermore, this telemedicine breast cancer diagnosis system will be designed using a web platform to help doctors easily access on different devices such as mobile phone or computer with the web. An important problem is that the computer application has good security and hardware support with a web

application which is compatible with a variety of the different devices, and is also easy to access anytime. Moreover, in this study, the choice of a web-based system is easy for remote access. Therefore, this system is designed with the following basic requirements: 1. Operate on a server in the Internet environment and provides access anytime, anywhere; 2. Support DICOM protocol and the ability to upload JPEG images; 3. Allow user communication based on a web platform; 4. Provide output results including lesion type, risk level of benign or malignant.

2.1. Hardware architecture of the telemedicine system

Figure 1 shows a proposed telemedicine system for remote breast cancer image classification. This system encompasses the core components, interacting peripheral devices, and the connections between them. In particular, the system is composed of three main components: 1. Picture Archiving and Communication System server (PACS) for images; 2. Web server for data connection with the AI model and PACS; 3. AI model for image classification.

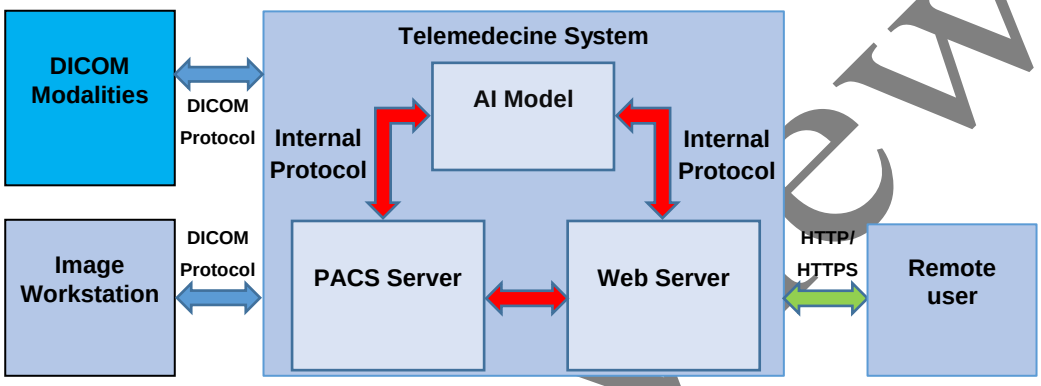


Figure 1. *Representation of the telemedicine system for classification of breast cancer images using a AI model.*

Therefore, this telemedicine system connects to peripheral devices such as X-ray imaging equipment, image viewing devices, and terminal devices for remote access. Details of each component in the system are described as follows:

(i) PACS server can receive and store images, particularly the images are uploaded from medical imaging devices with DICOM standard. Moreover, the system allows uploading the images in JPEG format for storing and processing. This server also interfaces with terminals, other components in the system through protocols and APIs.

(ii) Web server is to provide web services for creating an access interface for users. In particular, the web interface is built using modern web technology, providing high adaptability to many different types of terminal devices.

(iii) AI model is applied an EfficientNet-B7 model trained for classifying mammogram images. In particular, when receiving a request from the web server, the AI model performs the classification process and then returns classified results through the system's internal protocols.

(iv) Peripheral devices connected to the above system include: DICOM Modalities being breast X-ray imaging devices compatible with the DICOM standard; Imaging workstations for doctors to view images; remote users such as doctors or patients can access remotely via the website.

2.2. Description of EfficientNet-B7 model

In this system, the AI model plays an important role in the telemedicine diagnosis system of breast cancer problems. In particular, EfficientNet-B7 is used in the AI model for classifying 3 types of breast cancer lesions such as Benign, Malignant and Normal. Moreover, EfficientNet-B7 is one of the models suitable for breast cancer image classification. In particular, the Fully Connected (FC) and Classifier layers at the output of the pre-trained models are tuned to the breast cancer image sets for better classification performance. In this EfficientNet-B7 model, a Transfer Learning (TL) method is used with the Efficientnet-B7 pre-trained on ImageNet. In addition, EfficientNet-B7 features a compound scaling

method, which allows the depth, width, and resolution of the network to be adjusted simultaneously. Thus, the compound scaling formula is defined as follows:

$$d = \alpha^{\phi}, w = \beta^{\phi}, r = \gamma^{\phi} \quad (1)$$

in which α, β, γ are the scaling factors for depth, width, and resolution, respectively and ϕ is the compound coefficient.

EfficientNet-B7 is designed with scaling factors of α, β, γ determined through a neural architecture search and their following values are α=1.2, β=1.1, γ=1.15. In addition, EfficientNet-B7 uses a baseline network with the following parameters: input resolution=224×224, width=1, depth=1.

For high performance of classifying three types of breast lesions such as Benign, Malignant and Normal, we changed the final FC layer of the pre-trained EfficientNet-B7 model to an output layer with three units. In particular, the original FC layer was replaced by a subnetwork consisting of two Dense layers with a ReLU activation layer. In addition to this, determining the hyper-parameters for the pre-trained model is an important contribution. This process includes fine-tuning parameters such as batch size, number of epochs, optimizer, and learning rate to suit the training process:

- Batch size = 32: Each training run will process 32 data samples.
- Epochs = 100: The entire dataset will be passed through the model 100 times. The training process will end after completing 100 epochs.
- Optimizer = Adam: The Adam optimizer is chosen for its good adaptability and high performance in deep learning problems.
- Learning rate = 0.001: This learning rate determines the degree to which the model's weights are adjusted in each training step.

After training the model, we will evaluate its performance and accuracy to ensure its reliability. Evaluation criteria during training includes Loss Function (LF) using the CrossEntropyLoss function; a popular choice for classification problems. Thus, this function, which measures the difference between the model's predictions and the actual labels, is expressed as follows:

$$loss(x, y) = -\sum_{i=1}^{C} x_i \log(y_i) \quad (2)$$

in which $loss(x,y)$ is the value of the loss function, representing the error between the predicted value and the actual value. C is the number of classes. x_i is the actual label of the i^{th} class. If class i is the actual class of the data sample, $x_i = 1$, otherwise $x_i = 0$. y_i is the prediction probability of the model for the i^{th} class.

Accuracy is calculated as the percentage of correct predictions over the total number of data samples. In each epoch, accuracy is calculated for both the training dataset and the testing dataset.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

in which TP: Number of correctly predicted positive samples. TN: Number of correctly predicted negative samples. FP: Number of incorrectly predicted negative samples. FN: Number of incorrectly predicted positive samples.

2.3. Image processing

For better classification performance, the breast cancer images need to be preprocessed for extracting regions containing many features, as well as to synchronize the image size as described in Figure 2. Furthermore, the original image of the system is in DICOM format and has a large size and this is not suitable for the input of the AI classifier. Moreover, the original image contains some unnecessary background image information that may affect the classification accuracy. Therefore, image preprocessing is necessary before being fed into the system with the AI model as described in Figure 2.

In Figure 2, before being classified by the EfficientNet-B7 model, the images from PACS are processed through 3 steps: First, the original image in DICOM format is converted to JPEG format for matching the input format of the EfficientNet-B7 and the JPEG image still keep its unchanged size;

Next, the JPEG image is automatically cropped to focus on the part of breast object and also remove unnecessary components such as patient information and image background; Finally, the image is resized to be suitable to the input of the AI classifier.

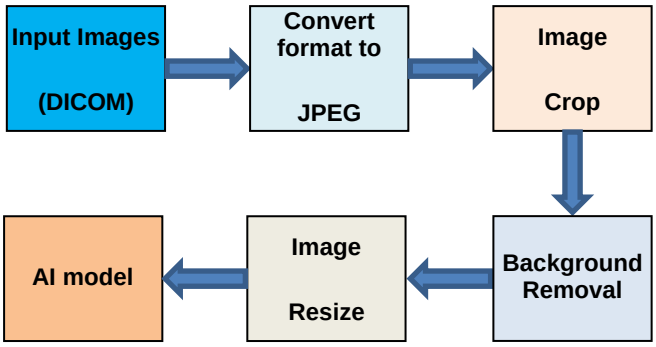


Figure 2. *Representation of Image pre-processing*

2.4. Describing internal protocols of the system core

Protocols in the telemedicine system play an important role. Therefore, in this system, we describe and propose some methods for the protocols to be compatible with the DICOM imaging system. In order for the parts of the system as described in Figure 1 to interact and exchange information each other, internal protocols are needed to perform this function. In the proposed system, the internal protocols, managing the communication between the PACS server, the web interface server and the EfficientNet-B7 model are the important part of the telemedicine system for breast cancer diagnosis. In particular, the PACS server, which stores medical images, relies on protocols to securely transmit relevant breast imaging studies to the web interface server upon user request. Therefore, the web interface server uses its own protocols to forward these images to the AI model for analysis, assigning tasks and the expected output format. Finally, the protocols determine how the AI model communicates its diagnostic predictions back to the web interface server, which then presents these results to the remote healthcare doctors. In this system, the internal communication paths are standardized to ensure data security and this will increase processing efficiency. With this good communication, the workflow will be highly reliable throughout the image retrieval process, as well as AI-assisted diagnosis.

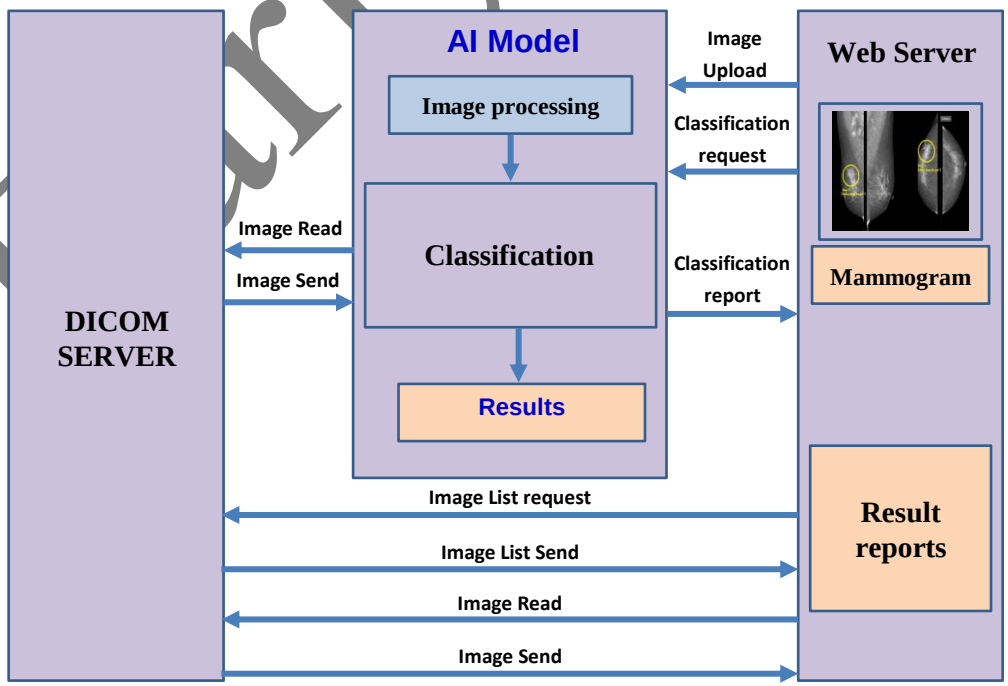


Figure 3. *Description of the internal protocols for information exchange between three main components in the system core.*

In this article, the AI model is designed to use the internal protocols for establishing to exchange information and coordinate with other components in the core of the system as depicted in Figure 2. In particular, the types of messages exchanged between the components need to be determined. In addition, the system is designed to interface with external peripherals such as X-ray devices, image display devices, and user terminals through appropriate protocols. Including the DICOM protocol for communication between the system and X-ray devices, DICOM image display devices already have the support of the protocol as a common standard in medicine. Image communication with user terminals uses the HTTP/HTTPS standard protocol for websites. The use of standard and popular protocols for external communications allows the system to be easily deployed and connected to existing medical facilities' equipment.

To process images before classification, the AI model needs to get images from the PACS server. Therefore, the communication between the AI model and the PACS needs to provide messages for performing this function. Meanwhile, the Web server needs not only to exchange images with the PACS server, but also to access the list of images stored on the server and other information which display on the web interface for users. Meanwhile, the communication between the web server and the AI model needs to provide a means to request the AI model for performing diagnosis and then returning results for generating user reports.

Furthermore, the proposed system is built using parameters for operating at the highest performance. Therefore, the specific system parameters relate to image size, storage capacity and other parameters which are determined based on the expected input data set, usage requirements and other system factors as described in Table 1.

Table 1. *Technical specifications of the telemedicine diagnostic system.*

No.	Parameters	Value
1	Storage capacity	200GByte
2	Transmission bandwidth	100Mbps
3	Maximum file size for upload	10Mbyte
4	Processing time	<20s

2.5. Data acquisition

In this study, DICOM mammogram images collected from the Oncology Hospital in Vietnam are used to train the AI model. For image collection, a strictly designed process is carried out to ensure the validity of the dataset as shown in Figure 3. First, the DICOM mammogram images are collected from the hospital's PACS image storage device and they are anonymized to protect patient privacy. Next, the dataset is annotated by some radiologists through a labelling tool on the hospital's computer. Finally, the annotated images are stored as a dataset.

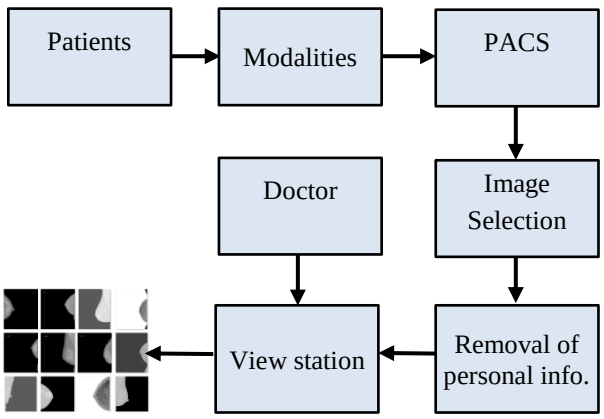
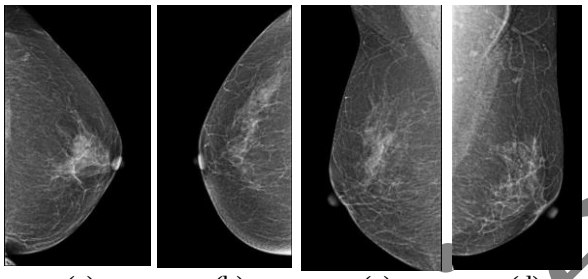


Figure 4. *Data collection diagram at the Ho Chi Minh City Oncology Hospital*

In this step, more than 4000 DICOM mammogram images from 2018 to 2023 in the PACS server of the Ho Chi Minh City Oncology Hospital DICOM system were selected. In particular, these images were collected on devices from the vendor SIEMENS. For ensuring patient privacy on the images such as identifiable patient information in the DICOM tags are completely erased. It means that only necessary information such as patient age and sex are retained. In addition to DICOM metadata, relevant information such as the angle and perspective of the image may appear in the images for preprocessing and feature extraction. Some sample images collected from the Ho Chi Minh City Oncology Hospital are shown in Figure 5. In addition, all are digital images collected from the DICOM system and unnecessary personal information has been removed. After preprocessing the images, the detailed information about the collected image set is described in Table 2.



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

Figure 5. *Sample images of a patient's exam at the Ho Chi Minh City Oncology Hospital: (a) Left CC; (b) Right CC; (c) Right MLO; (d) Left MLO*

Table 2. *Representation of the detail information of an image after preprocessing.*

No.	Parameters	Value
1	Number of Images	4034
2	Image size	1024×1024
3	Format	DICOM
4	File size	8.8MB

For the purpose of this study, the image collection was diagnosed and classified by specialists into three primary categories: normal (no lesions), benign, and malignant, as presented in Table 3. To ensure a balance in the number of images during training, images from the original image set were reselected with a number of 800 images per category.

Table 3. *Representation of the number of lesion types in the image set*

No.	Classification	Number of Images	Training	Testing
1	Normal	1815	640	160
2	Benign	877	640	160
3	Malignant	838	640	160

3. Results and Discussion

After completion of the telemedicine diagnosis system, the results are obtained for evaluating and we would like to show two main independent aspects: Classification accuracy of the EfficientNet-B7 model; Operational parameters of the telemedicine system. Therefore, each aspect is evaluated through many parameters for determining the effectiveness of the proposed methods, as well as the level of meeting the requirements. During this evaluation, the system indicators some problems, including accuracy, system reliability, transmission parameters, processing time, delay, protocols and other parameters.

3.1. Result of Breast cancer classification

In this article, the classification accuracy is evaluated through the confusion matrix, in which the image sets are collected from the Ho Chi Minh city Oncology Hospital and the test dataset is completely independent of the dataset for training, with a quantity ratio of 8:2. Basically, there are 3 types of image sets such as Benign, Malignant and Normal. Therefore, the training results of the EfficientNet-B7 model are evaluated by the Loss and Accuracy functions.

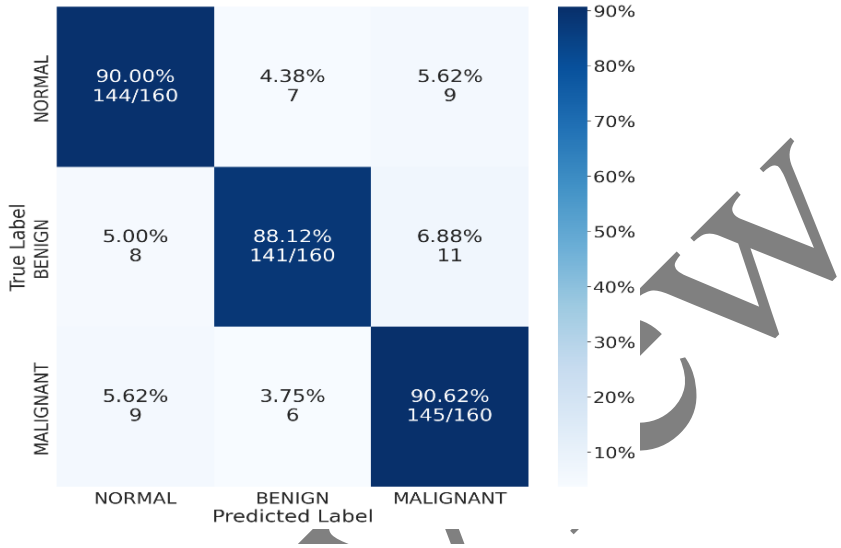


Figure 6. *Representation of the confusion matrix for three types of lesion such as Benign, Malignant and Normal.*

Figure 6 presents the confusion matrix, showing the model's classification ability for three breast lesion categories: Benign, Malignant and Normal. In particular, the EfficientNet-B7 model predicted related to correction and mistake as shown in Table 4. It means that when classifying, the category of Benign has 141 of True Positive and in case of False Positive, it recognizes 11 Benign to Malignant and 8 Benign to Normal, it is similar to the categories of Malignant and Normal.

Table 4. *Representation of the prediction related to the classification using the confusion matrix.*

Categories	Number of Images	True Positive	False Positive	
Normal	160	144	7 to Benign	9 to Malignant
Benign	160	141	8 to Normal	11 to Malignant
Malignant	160	145	9 to Normal	6 to Benign

Analysis of the training results shows that the AI model with EfficientNet-B7 achieves high training accuracy, particularly converging at around 97%, indicating good learning ability on the training data. However, the accuracy on the test set is only around 89.58 % and this may be a small problem related to constructing new data. The difference between the training and testing accuracy may be due to the fact that the data set used is real data, has a large difference in quality and is collected from different devices. The precision and F1-score are showed in Table 5.

Table 5. *Representation of the precision and F1-score.*

Categories	Precision	recall	f1-score
Normal	0.8944	0.9000	0.8972
Benign	0.9156	0.8812	0.8981
Malignant	0.8788	0.9062	0.8923

3.2. Result of Telemedicine model

The system is evaluated for telemedicine functionality through comprehensive tests of data transmission, stability, and system bandwidth through the following detailed test steps.

- Simulated data transmission: Simulate the transmission of a diverse set of test images through the system, simulating various network conditions such as different bandwidths or latency.
- Latency measurement: Measure the time required for images to be uploaded, processed by the AI model, and displayed; Record these delays.
- Error rate assessment: Introduce controlled errors during the simulated transmission such as packet loss to evaluate the system's robustness and error handling; monitor for data corruption or system failures.
- System uptime monitoring: monitor the uptime and stability of the web interface server and AI model over a period of time and record any errors.
- Qualitative evaluation: Evaluate the user experience of the web interface in terms of ease of use and clarity of presentation results.

The evaluation steps provided data for calculating specific parameters as follows: Average latency for processing images and delivering results; Transmission error rate or system failure rate under stress; System uptime rate. In addition, this protocol provides a framework for evaluating a telemedicine system for breast cancer diagnosis based on EfficientNet-B7. Thus, specific parameters and thresholds for acceptable performance will be determined based on the requirements and context of the intended application. During the evaluation of the imaging data, the real DICOM data are collected.

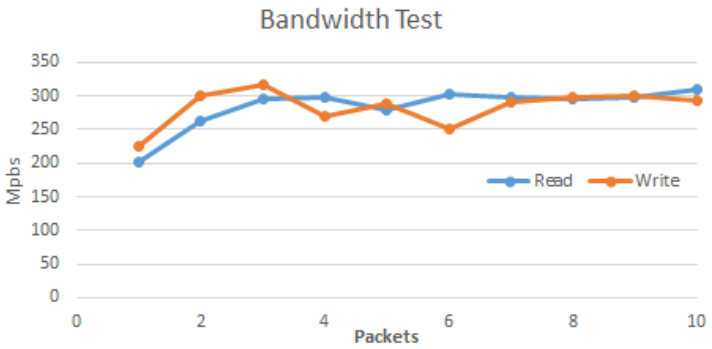


Figure 7. *Measurement results of the system bandwidth.*

Figure 7 shows the results of testing the telemedicine system. In particular, the graph shows that the system Read/Write speed is always above the allowable limit of 200Mbps and the average speed is 284Mbps.

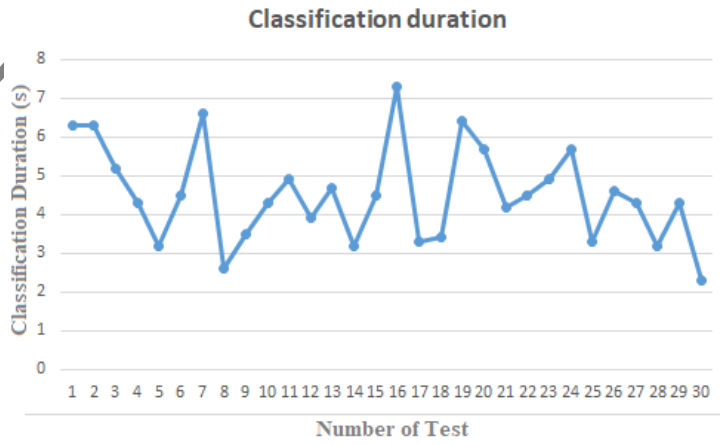


Figure 8. *Measurement results of the classification duration.*

Figure 8 shows the test results of the system's classification time with 30 trials. In particular, the results show that the longest classification time is 7.3s and the fastest is 2.3s and the average classification time is 4.514s.

Table 6. *Representation of the system evaluation results.*

Parameter	Number of Test	Result	Standard deviation
Accuracy	1500	84%	-
Bandwidth	-	284Mbps	32.23
Fault ratio	1500	0.47%	-
Average classification duration	30	4.514s	1.52

The results are presented in Table 6, in which the system meets the requirements of a system for practical remote diagnosis. However, to achieve higher performance and accuracy, some techniques need to be followed such as: input image quality related to substandard lighting, noisy images; The requirement of large computational resources for the training process is a barrier for units with limited infrastructure.

This result is comparable to other telemedicine systems for breast cancer diagnosis. For example, VinDr AI's solution for breast cancer diagnosis - VinDr-Mammo system has a practical BI-RADS classification accuracy of over 80% [19]. Other technical parameters of similar systems are not published, making comparison difficult.

3.3. Discussion

In recent years, deep learning networks have attracted researchers for image classification in different fields, in which medical images for diagnosis support are classified. A study utilizing YOLOv6 combined with Federated Learning (FedL) on the BreakHis and BUSI datasets achieved a remarkable 98% accuracy in differentiating between benign and malignant tissues [10]. A research is that general an AI-based diagnostic systems produced a report to reach an accuracy as high as 94.5% [14]. Although, our results of the proposed telemedicine system have not yet reached this state-of-the-art threshold, its strength points are to complete the telemedicine system architecture from input data to end results. In particular, the performance of our proposed AI model presents a nuanced picture and the 89.58% accuracy is a significant step towards automated diagnosis. In addition, our work addresses the practical challenge of integrating the model within a telemedicine infrastructure, including PACS server communication, web-based user access, and defined internal protocols for data exchange. Focusing on telemedicine the system for application is a key contribution in this article.

Moreover, our research aligns with developing the telemedicine system using AI model in medicine area for reducing diagnostic errors and physician workload. Although, we did not directly measure error reduction using the telemedicine compared to doctor diagnosis, the high accuracy in identifying the Normal cases of 92% precision as shown in Figure 5 is very significant. Thus, it means that our proposed system could effectively produce a significant number of right cases and this can allow specialists to focus on more ambiguous or potentially malignant ones. In addition, it aligns with the 44% reduction in physician workload reported by MammoScreen [20].

In this article, the classification accuracy of the EfficientNet-B7 model in the telemedicine system is not the highest compared to previous researches, but our telemedicine system has made a significant contribution by presenting a complete, practical, and efficient telemedicine system designed for real-world deployment. In particular, the system's architecture is robust and its performance can allow to apply for clinical testing. Therefore, this system can focus on enhancing and improving its diagnostic accuracy to be able to apple at Vietnam hospitals.

4. Conclusions

This article proposes a telemedicine system for classifying breast cancer images using an EfficientNet-B7 model and this system can be applied for classifying breast lesions at hospitals in

Vietnam. In the proposed telemedicine system with the EfficientNet-B7 model, a dataset of DICOM mammogram images from the Oncology Hospital in Ho Chi Minh City were used to train this EfficientNet-B7. The results produced a high training accuracy of approximately 97%. However, the validation accuracy on the test set was just 89.58%.

The importance is that the telemedicine system demonstrated robust performance, meeting the technical requirements for practical deployment. In particular, it achieved an average data transmission speed of 284Mbps, which is higher than the 200Mbps requirement and also an average classification time of 4.514 seconds per image. These results illustrate the system's capability to operate efficiently for real clinical applications and this can provide timely support for many doctors in remote diagnosis at the same time. In addition to this, our telemedicine system shows great promise in reducing physician workload and improving diagnostic access. This system can focus on enhancing the model's accuracy, in which image sets should be obtained more and AI model should be improved and upgraded.

Acknowledgments

We would like to thank the Ho Chi Minh City University of Technology and Education (HCMUTE), Vietnam.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] World Health Organization, "Vietnam source: Globocan 2022," 2022. [Online]. Available: https://gco.iarc.who.int/media/globocan/factsheets/populations/704-viet-nam-fact-sheet.pdf
[2] Vietnamnews, "Bach Mai Hospital leads digital transformation in healthcare," *Vietnamnews*, 2025. [Online]. Available: https://vietnamnews.vn/opinion/1693875/bach-mai-hospital-leads-digital-transformation-in-healthcare.html. Accessed: Jun. 18, 2025.
[3] K. N. Ramanto and A. A. Parikesit, "The usage of deep learning algorithm in medical diagnostic of breast cancer," *Malays. J. Fundam. Appl. Sci.*, vol. 15, 2019, doi: 10.11113/mjfas.v15n2.1231.
[4] Y. Deng *et al.*, "A new framework to reduce doctor's workload for medical image annotation," *IEEE Access*, vol. 7, pp. 107097–107104, 2019, doi: 10.1109/ACCESS.2019.2917932.
[5] N. N. Trung, "Updating and upgrading the features of artificial intelligence (AI) applications in early diagnosis of breast cancer with Korean experts at Thai Binh University of Medicine and Pharmacy Hospital," *Thai Binh Univ. Med. Pharm. Web Portal*, 2025. [Online]. Available: https://tbump.edu.vn/tin-hoat-dong/cap-nhat-nang-cap-tinh-nang-cua-ung-dung-tri-tue-nhan-tao-ai-trong-chan-doan-som-ung-thu-vu-cung-cac-chuyen-gia-han-quoc-tai-benh-vien-dai-hoc-y-thai-binh-298.html. Accessed: Jun. 16, 2025.
[6] S. A. A. Karim, U. H. Mohamad, and N. E. N. Puteri, "Discovery of interpretable patterns of breast cancer diagnosis via class association rule mining (CARM) with SHAP-based explainable AI (XAI)," *Malays. J. Fundam. Appl. Sci.*, vol. 21, 2025, doi: 10.11113/mjfas.v21n3.3792.
[7] J. Yan *et al.*, "Diagnosis and treatment of breast cancer in the precision medicine era," in *Precision Medicine*, T. Huang, Ed. New York, NY, USA: Springer, 2020, pp. 53–61, doi: 10.1007/978-1-0716-0904-0_5.
[8] K. K. Evans, R. L. Birdwell, and J. M. Wolfe, "If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening," *PLoS One*, vol. 8, 2013, Art. no. e64366, doi: 10.1371/journal.pone.0064366.
[9] M. Issaiy, D. Zarei, and A. Saghazadeh, "Artificial intelligence and acute appendicitis: A systematic review of diagnostic and prognostic models," *World J. Emerg. Surg.*, vol. 18, 2023, Art. no. 59, doi: 10.1186/s13017-023-00527-2.
[10] C. Gupta *et al.*, "Applying YOLOv6 as an ensemble federated learning framework to classify breast cancer pathology images," *Sci. Rep.*, vol. 15, 2025, Art. no. 3769, doi: 10.1038/s41598-024-80187-7.
[11] A. Kapoor, P. Nambisan, and E. Baker, "Mobile applications for breast cancer survivorship and self-management: A systematic review," *Health Informatics J.*, vol. 26, pp. 2892–2905, 2020, doi: 10.1177/1460458220950853.
[12] J. V. Johansson and E. Engström, "'Humans think outside the pixels' – Radiologists' perceptions of using artificial intelligence for breast cancer detection in mammography screening in a clinical setting," *Health Informatics J.*, vol. 30, 2024, Art. no. 14604582241275020, doi: 10.1177/14604582241275020.
[13] N. Zhou *et al.*, "Concordance study between IBM Watson for Oncology and clinical practice for patients with cancer in China," *Oncologist*, vol. 24, pp. 812–819, 2019, doi: 10.1634/theoncologist.2018-0255.
[14] Y. Shen *et al.*, "Artificial intelligence system reduces false-positive findings in the interpretation of breast ultrasound exams," *MedRxiv*, 2021, doi: 10.1101/2021.04.28.21256203.
[15] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," 2020. [Online]. Available: https://arxiv.org/abs/1905.11946
[16] R. Raza *et al.*, "Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images," *Eng. Appl. Artif. Intell.*, vol. 126, 2023, Art. no. 106902, doi: 10.1016/j.engappai.2023.106902.
[17] A. Abdelrahman and S. Viriri, "EfficientNet family U-Net models for deep learning semantic segmentation of kidney tumors on CT images," *Front. Comput. Sci.*, vol. 5, 2023, doi: 10.3389/fcomp.2023.1235622.
[18] Y. Sun, Z. Zhu, and B. Honarvar Shakibaei Asli, "Automated classification and segmentation and feature extraction from breast imaging data," *Electronics*, vol. 13, 2024, doi: 10.3390/electronics13193814.

[19] VinDr, "VinDr – Solutions for medical data," [Online]. Available: https://vindr.ai/. Accessed: Aug. 1, 2025.
[20] K. Lang *et al.*, "Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded screening," *Lancet Oncol.*, vol. 24, pp. 936–944, 2023, doi: 10.1016/S1470-2045(23)00298-X.

Thanh-Tam Nguyen is a lecturer at School of Biomedical Engineering, International University Vietnam National University - Ho Chi Minh City, Vietnam. He received the B.E. and the M.E degrees in Electronic and Telecommunication Engineering from Ho Chi Minh City University of Technology, Viet Nam, in 2002 and 2004, respectively. He currently working toward the Ph.D. degree in Electronic Engineering at Ho Chi Minh City University of Technology and Education, Viet Nam.
His main research interests include Telemedicine, Machine learning, Telecommunication, Medical Instrumentation.
Email: tamnt.ncs@hcmute.edu.vn. ORCID: https://orcid.org/0000-0002-7920-1651

Thanh-Hai Nguyen is a lecturer and Head of Department of Industrial Electronics and Biomedical Engineering, Faculty of Electrical – Electronics Engineering, HCMC University of Technology and Education, Ho Chi Minh City, Vietnam. He received the B.S. degree in Engineering of Electrical and Electronics from Ho Chi Minh City University of Technology and Education, Vietnam, and the M.E, degree in Electronic-Telecommunication from Ho Chi Minh City University of Technology, Vietnam. He obtained his Ph.D. in Electrical and Electronics Engineering from University of Technology, Sydney, Australia. His main research interests include Signal-Image processing, Biomedical Engineering, Smart wheelchairs, Machine Learning, Artificial Intelligence.
Email: nthai@hcmute.edu.vn. ORCID: https://orcid.org/0000-0003-3270-6975

Tin-Trung Nguyen is Head of Department of Radiology, Ho Chi Minh City Oncology Hospital, Ho Chi Minh City, Vietnam.
His main research interests include Telemedicine, Breast cancer, Lung cancer.
Email: nguyentintrung.dr@gmail.com. ORCID: https://orcid.org/0000-0002-8894-4289