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.

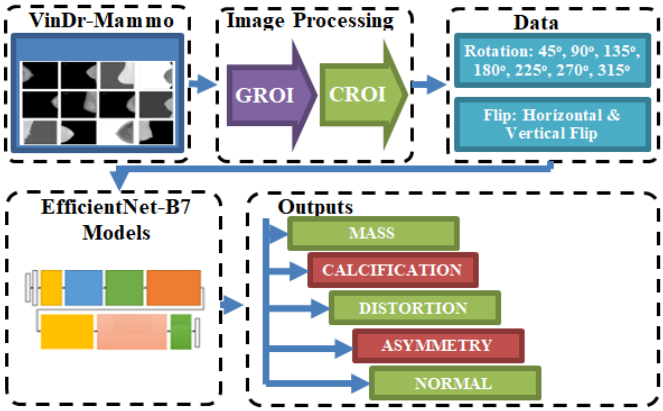


Fig. 1. Block diagram of the proposed method for breast lesion classification.

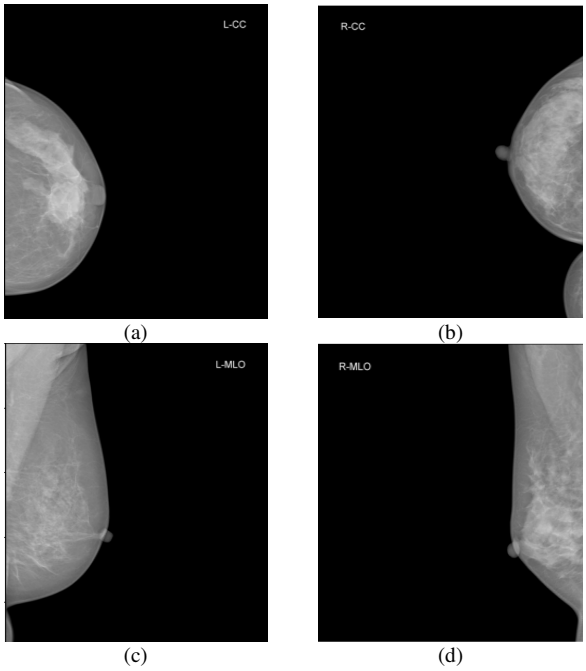


Fig. 2. A sample set of four breast images from a single mammography exam in the VinDr-Mammo dataset: (a) left CC, (b) right CC, (c) left MLO, (d) right MLO.

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.

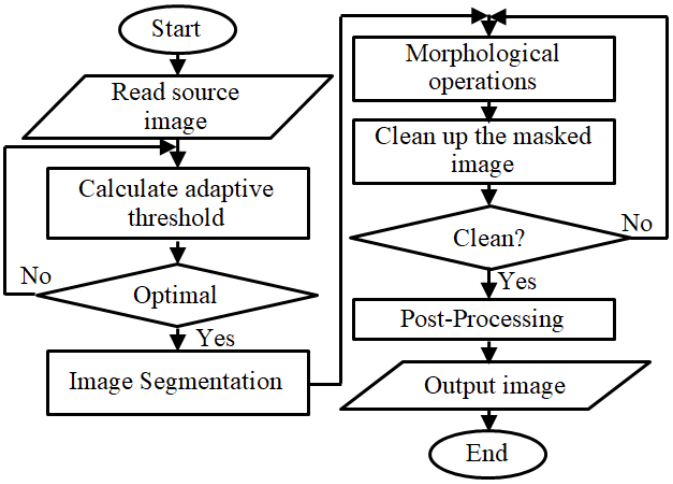


Fig. 3. Workflow of Algorithm 1 for determining the Otsu threshold to extract the GROI from mammography images.

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.

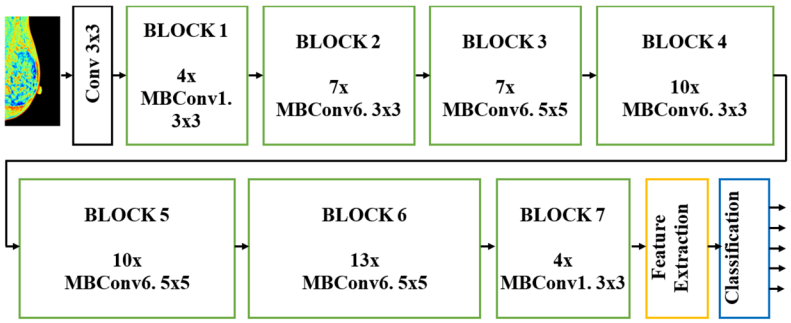


Fig. 4. Architecture of the fine-tuned EfficientNet-B7 for classifying five categories of mammography image sets.

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

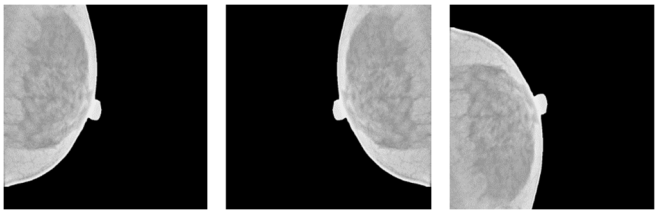
A. Augmentation of Image Sets

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.

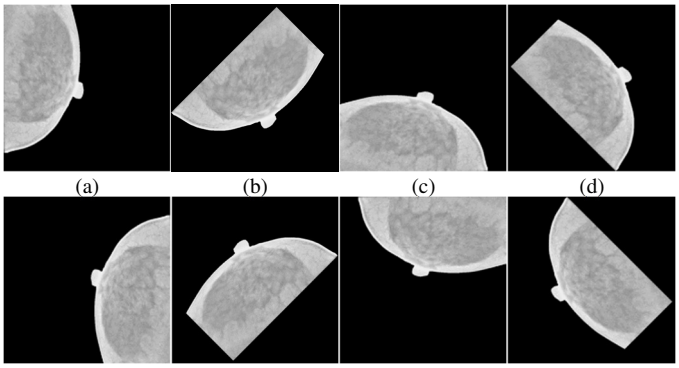
TABLE II. ORIGINAL IMAGE COUNTS, TRAIN/TEST SPLITS, AND BALANCED DATASET SIZES AFTER AUGMENTATION OR REDUCTION

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°.

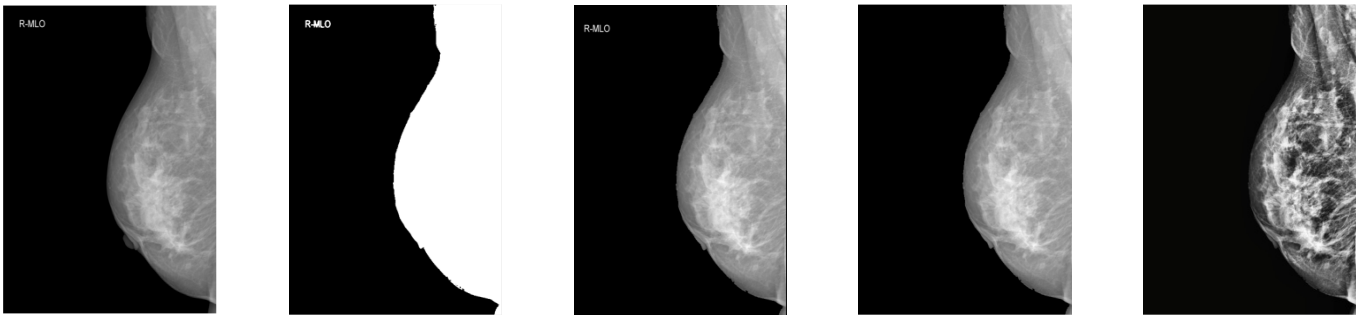
B. Results of GROI and CROI Extraction

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.

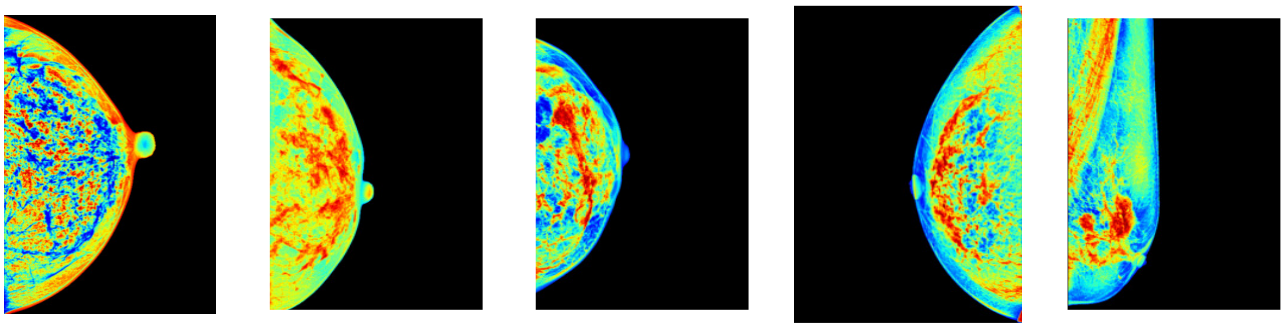


(a) (b) (c) (d) (e)

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)

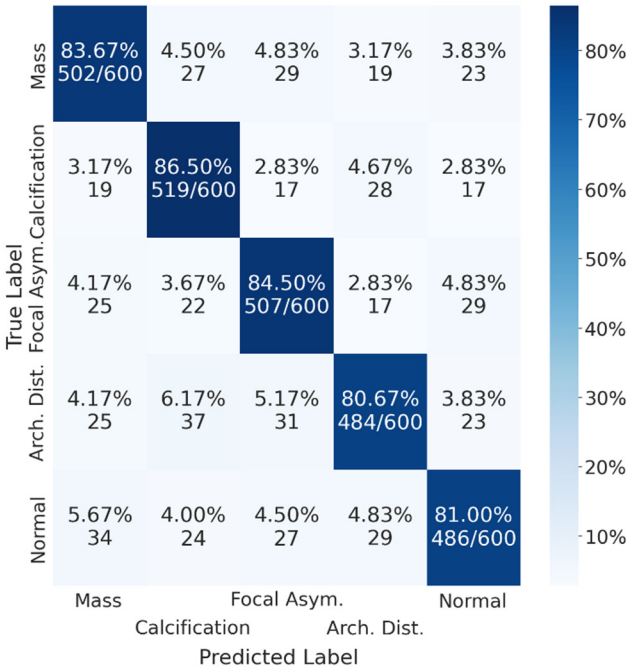
Fig. 9. Representation of five images categories with CROI: (a) normal, (b) focal asymmetry, (c) architectural distortion, (d) calcification, (e) mass.

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.

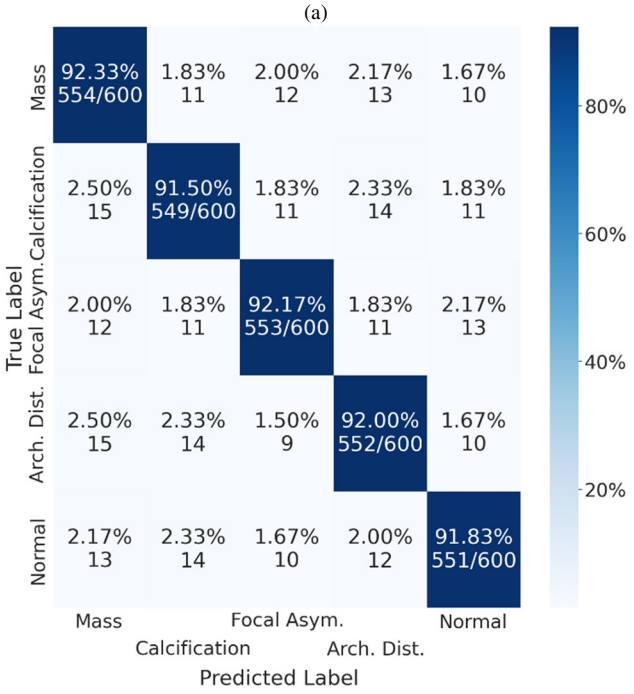
C. Comparison with Related Work

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.



(a)



(b)

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.