EMITTER International Journal of Engineering Technology
Vol. 10, No. 1, June 2022, pp. 136~152
DOI : 10.24003/emitter.v10i1.695

An Image Processing Framework for Breast Cancer Detection Using Multi-View Mammographic Images

Nada Fitrieyatul Hikmah[1], Tri Arief Sardjono[1], Windy Deftia Mertiana[1], Nabila Puspita Firdi[1], Diana Purwitasari[2]

[1]Department of Biomedical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2]Department of Informatics Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
Correspondence Author: nadafh@bme.its.ac.id

Received March 6, 2022; Revised April 7, 2022; Accepted May 9, 2022

Abstract

Breast cancer is the leading cause of cancer death in women. The early phase of breast cancer is asymptomatic, without any signs or symptoms. The earlier breast cancer can be detected, the greater chance of cure. Early detection using screening mammography is a common step for detecting the presence of breast cancer. Many studies of computer-based using breast cancer detection have been done previously. However, the detection process for craniocaudal (CC) view and mediolateral oblique (MLO) view angles were done separately. This study aims to improve the detection performance for breast cancer diagnosis with CC and MLO view analysis. An image processing framework for multi-view screening was used to improve the diagnostic results rather than single-view. Image enhancement, segmentation, and feature extraction are all part of the framework provided in this study. The stages of image quality improvement are very important because the contrast of mammographic images is relatively low, so it often overlaps between cancer tissue and normal tissue. Texture-based segmentation utilizing the first-order local entropy approach was used to segment the images. The value of the radius and the region of probable cancer were calculated using the findings of feature extraction. The results of this study show the accuracy of breast cancer detection using CC and MLO views were 88.0% and 80.5% respectively. The proposed framework was useful in the diagnosis of breast cancer, that the detection results and features help clinicians in making treatment.

Keywords: Breast Cancer, CC view, Entropy, Feature Extraction, Mammography, MLO view.

1. INTRODUCTION

Cancer is a leading cause of worldwide's death. Global Burden of Cancer Study (Globocan) from World Health Organization (WHO) estimated 19.3 million new cancer cases and almost 10.0 million cancer deaths occurred in

2020 globally. Female breast cancer is the most commonly diagnosed cancer and the leading cause of cancer death with an estimated 2.3 million new cases in 2020. It represents 11.7% of all cancer cases [1]. According to data released by Globocan in 2020, breast cancer has the highest number of new cases in Indonesia with 65,858 cases or 16.6% of the total 396,914 cancer cases.

The breast consists of glandular tissue that includes mammary gland tissue, fat, and connective tissue. If the cells in the mammary glands divide and grow abnormally, these cells can develop into benign or malignant tumors [2]. Women have a higher rate of breast cancer than men. Factors that influence breast cancer include age, family history or genetics, menstrual cycle, diet, lifestyle, childbirth, drug use, and history of cancer [3].

The early phase of breast cancer is asymptomatic, without any signs or symptoms. The presence of a lump or thickening in the breast is the most common sign and symptom of breast cancer. The earlier breast cancer can be detected, the treatment will be easier and the chance of cure is greater [4][5][6][29]. Early detection can be physical and supporting examinations. Early detection begins with breast self-examination to detect an abnormal lump in the breast, a change in skin color, or a change in the nipple. If an abnormality is found on the breast self-examination, the next examinations are clinical examination and supporting examinations such as screening mammography. The role of radiological imaging is very important in cases of breast cancer, both palpable and non-palpable [7]. When screening for breast cancer, there are a few different types of radiological examinations that can be performed. Mammography, ultrasonography, and magnetic resonance imaging are the three modalities of radiology that are used the most commonly [2][3].

Examination with the screening mammography is considered the most effective for detecting the possibility of cancer tissue in the breast [8]. Lumps of breast cancer often appear on mammographic images even though they are typically low contrast and blurry [9]. Mammography is also inexpensive and has a high sensitivity for lumps and microcalcifications [10]. Although breast cancer can be diagnosed through mammographic images by radiologists, there is a possibility of inaccuracy considering human visual limitation and objectivity. For this reason, computer-based systems that can assist the diagnosis process have been developed in recent years [9][10].

According to Sickles et al. [11], there are five stages of radiologists in assessing mammography results, which are indications, analyzing the category of breast tissue, analyzing important findings, comparing with previous studies, and determining the final assessment according to the Breast Imaging Reporting and Data System category (BI-RADS). In mammography, two angle views are used to evaluate abnormalities in the patient's breast. Craniocaudal projection, often known as an overhead view, and mediolateral oblique, sometimes known as a side view [2] are the two perspectives that are available. The medial and exterior lateral portions of the breast are shown in the craniocaudal projection, which is taken from the top to the bottom of the breast. Meanwhile, the mediolateral oblique projection aims to visualize the entire breast taken from an angle between 40^0 and 60^0 from side. MLO view usually shows the lymph nodes with pectoral muscles.

Computer-Aided Diagnosis (CAD) for breast imaging plays a significant role in breast cancer screening, detection, and identification. The CAD system aids radiologists in the early detection of breast cancer. Numerous studies have utilized the CAD approach to classify mammograms using distinct MLO or CC images. The diagnostic performance of multi-view screening is superior than that of single-view screening. If there are more mammographic images, it is easier to identify and diagnose breast cancer [12]. This is applicable for both early detection and advanced stages of the disease. The combination of mammography image analysis methods that can detect breast cancer in CC view and MLO view is crucial to assist clinicians in determining breast cancer.

This paper focuses on automated breast cancer detection of mammographic images based on CC and MLO views. Our main goal is to develop an image processing framework to detect breast cancer using multi-view mammographic images. In addition, we propose enhancement, segmentation, and feature extraction of mammographic images. Segmentation produces the region of interest (ROI) from cancer for further calculations to obtain radius and area information through the feature extraction process.

The rest of this paper is organized as follows. Part 2 reviews existing works on breast cancer detection and discuss their limitations. Part 3 describes the originality and the uniqueness of the proposed method. Part 4 presents details of our image processing framework and region of interest estimation methods. Results and discussion are provided in Part 5. Part 6 concludes the research.

2. RELATED WORKS

Recently, few research efforts automated the detection of breast cancer. Sama et al. [5] developed a breast cancer classification method based on the entropy method applied to the Digital Database for Screening Mammography (DDSM) data. The entropy method was used in conjunction with the evolutionary algorithm in the mass phase to improve the performance of the evolutionary algorithm as feature extraction. The results of feature extraction were used as input to the fuzzy method to classify benign and malignant cancers. This study had sensitivity and specificity values of 72.88% and 64%, respectively.

Sepulveda et al. [13] proposed the detection of breast cancer using the local entropy method. In this study, pre-processing was applied using a high pass filter. Segmentation was carried out using the local entropy method. In this study, a morphological operator was applied as noise removal on binary images obtained at the thresholding stage. However, this study has not explained the optimum threshold value that can be applied to the entire image. The algorithm in each experiment still used different threshold values. Singh et al. [9] developed the detection and segmentation of breast cancer using a statistical approach. This study used intensity to represent the cancer area. The max-mean method and the least variance technique were applied to certain windows to identify the cancer area. In this study, a morphological operator was applied as a region patch and an image gradient technique to obtain the edges of the cancer area.

Simple threshold-based techniques can be used to segment regions with similar pixel intensities. Pectoral muscle segmentation using histogram-based thresholding requires a combination of morphological operations to identify muscle regions. Global threshold with morphological operation method proposed by Unni et al. [14] to estimate the pectoral muscle line. A similar approach to estimating the pectoral muscle line was also carried out by [15]. Those studies show the results of segmentation images without providing accuracy validation.

The next research proposed by [16], detected the pectoral muscle using a geometry-based method. This method succeeded in overcoming the shortcomings that occurred in the research conducted by [17]. Despite the existence of a convoluted density structure and changing curvature of the pectoral muscles, this study was effective in detecting the pectoral muscle in mammographic images. The researchers were able to identify and extract the contours of the breast by using canny edge detection. This approach achieved 95% accuracy on 322 Mammographic Image Analysis Society (MIAS) images.

The Hough Transform was employed in the study that was referenced in [18], and it was applied to 322 MIAS pictures in order to determine the border of the pectoral muscle line. The researchers were able to achieve an accuracy of 97.08%. In research [19], breast cancer images were preprocessed using Hough Transform and Canny Edge Detection to build a dataset for Deep-CNN, with a pectoral muscle removal accuracy of 99.06%. The Hough transform is an excellent tool for determining pectoral muscle borders, according to this study. This approach can be used to identify existing cancer lumps and determine their radius.

Based on the previous works reviewed above, the research on breast cancer detection has been carried out a lot. However, there is no additional information about radius and ROI that is used for diagnosis reference [5][8][9][13][20]. Therefore, this paper proposes a computer-based system based on CC and MLO views that contains contrast enhancement, segmentation, and feature extraction method with the additional information on radius and area of detected breast cancer. Indirect contrast enhancement becomes the most used method in mammographic images processings. This method does not directly change the structure of the image but through histogram manipulations. In segmentations, a texture-based analysis is used to differentiate between the object as foreground and surrounded tissue as background. Then, feature extraction is performed to get the radius and area information, which is expected to assist the clinicians throughout the diagnosis process so they can decide the treatment properly.

3. ORIGINALITY

Screening mammography is widely chosen as one of the modalities for imaging the breast. Clumps and microcalcifications appear on mammographic images with typical low contrast, making them difficult to detect. Therefore, the image enhancement in mammographic images is to increase the contrast so that it can distinguish objects and backgrounds with visual inspection [21]. Contrast enhancement can be done in a direct and indirect method. The indirect method commonly used is histogram equalization. However, histogram equalization significantly changes the brightness of an image so that the image looks unreal [22]. In histogram equalization, the brightness of an image is always considered at the middle gray level which cannot be applied to all input images, because brightness preservation is an important aspect of an image. Contrast Limited Adaptive Histogram Equalization (CLAHE) is a suitable method for sharpening the contrast of mammographic images [21]. In addition, the CLAHE method can be applied to mammographic images with dense breast tissue characteristics [23].

In the lumps and microcalcifications detection of breast cancer, a segmentation process is needed to separate the part that is considered cancer tissue from the surrounding tissue. The scheme is carried out by extracting ROI which is estimated to contain clumps [13]. The process can be done by using the texture analysis method. It analyzes the grayscale change in the image which contains important information about the transition area. One of the statistical measurements of randomness for characterizing the texture of an image is Entropy which indicates the complexity of the image. A higher entropy value indicates a more complex image. For this reason, the entropy-based algorithm is suitable to represent the transition region.

Following the detection of a breast cancer lump, feature extraction is conducted to acquire the segmentation radius and region of probable cancer. In the feature extraction process, the border following method is used which can represent the topology of thresholding binary mammography images' results. The border following algorithm is useful for getting the relationship between the area around the border with the outer border and hole border conditions that have a one-to-one correspondence with connected components. The border following algorithm considers the area within the border as a cancer area [8].

Segmentation of the pectoral muscle from the MLO view is a major challenge. Many previous studies have focused on the removal of pectoral muscle only, and few studies have performed an overall parameters assessment of the two mammographic views. To realize the issues of automatic segmentation of the pectoral muscle and breast cancer, this study started by extracting the region using classical image processing. Briefly, several key contributions of this paper can be highlighted: (a) pectoral muscle removal from the MLO view, (b) the image processing framework to detect areas of suspected breast cancer based on multi-view mammographic images, and (c) the assessment of both CC and MLO views to get radius and ROI from suspected breast cancer. Our proposed method assesses feature extraction which is one of the benchmarks for the stage of breast cancer. In addition, this information can also help in deciding treatment for the patient.

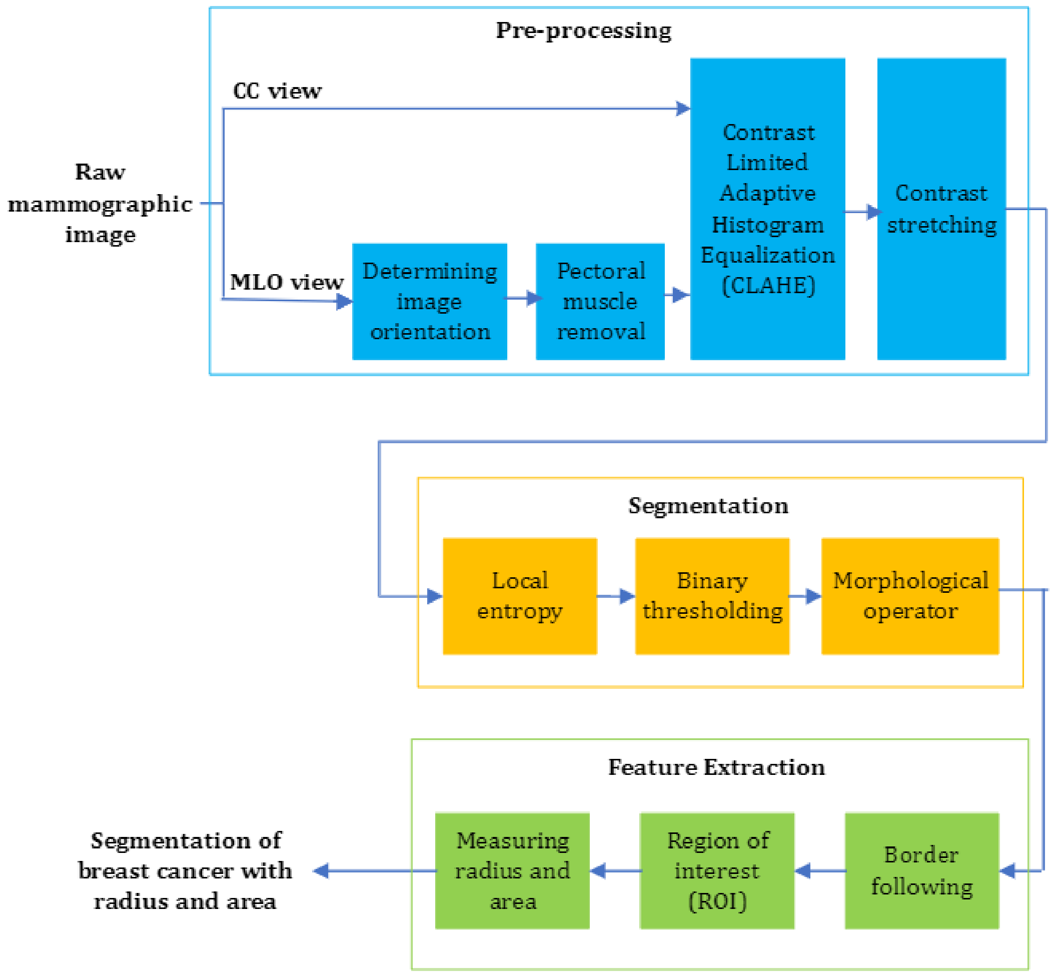

Figure 1. Framework for breast cancer detection.

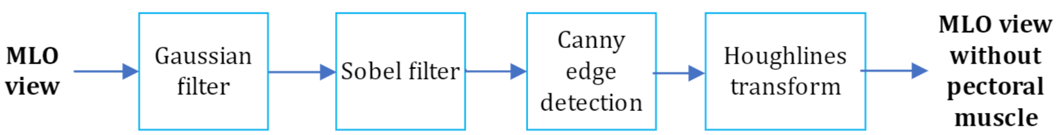

Figure 2. The stages of pectoral muscle removal.

4. SYSTEM DESIGN

As illustrated in Figure 1, the system architecture consisted of three primary phases: preprocessing, segmentation, and feature extraction. In preprocessing, raw mammographic images were compressed to minimize the memory used. There was a slight difference in the pre-processing stage between CC and MLO views. In the MLO view, the orientation of the image was determined and the pectoral muscles were removed, which is modified from the method according to [30]. In order to get the desired level of contrast, we first modified the adaptive histogram using Contrast Limited Adaptive Histogram Equalization (CLAHE), then we used the contrast stretching approach. Texture-based segmentation was applied to distinguish between an object as the foreground and the surrounding tissue as the background. Using the first-order local entropy, cancer-suspected tissue was obtained in this procedure. Based on segmentation results, the border-following algorithm was implemented. This method will generate edge points of the object which was used to locate ROI. Each stage of the system design will be explained in the following subsection.

4.1 Dataset

The online database from Digital Database for Screening Mammography (DDSM) [24] is used for this research. The research used 120 images of CC view and 60 images of MLO view. The raw images had a resolution of 500 microns. According to the BI-RADS evaluation for normal or cancerous breast tissue, the images were selected based on breast tissue types 1 (completely fatty), 2 (scattered fibroglandular density), and 3 (non - uniformly dense) [11].

4.2 Pre-processing

Raw mammographic images were compressed by decreasing the resolution from 500 microns to 50 microns using the downsampling method. The pre-processing stage in this research was divided into 2 different methods between CC and MLO views. Determination of image orientation was done in MLO view. A patient had 2 mammographic images for the left and right breast images. In order to facilitate the second step of processing, one of the mammographic images must be flipped so that they are both orientated in the same orientation. Figure 2 depicts the subsequent phase, the excision of the pectoral muscle. The initial step was to identify the pectoral muscle border line's outline. In order to get a crisp contour of the breast, a Gaussian filter and Sobel filter were used. The contours produced by the Sobel filter were transformed into lines by the application of edge detection. Hough lines transform approach was used to estimate the shape of the pectoral muscles.

The next pre-processing stage was applied to CC and MLO views. In contrast-enhancement, two stages of histogram modification were used. CLAHE was applied followed by contrast stretching for histogram normalization. This should be done because cancerous tissue on mammographic images often appeared in low contrast and overlapped between normal tissues. CLAHE is affected by the clip limit parameter which will affect the cropping level value in the image. The greater value of the clip limit causes the image to be over contrasted so that the image will be too bright and have an impact on the next process. This study used a clip limit of 2.0 and a structural element of 8x8.

After CLAHE, contrast stretching was applied. Contrast stretching was done by calculating the upper bound and lower bound of the gray level [25] as shown in Equation 1,

$$P_{out} = \left(P_m - c\right)\left(\frac{b-a}{d-c}\right) + a \tag{1}$$

where P_m is the input image, a is the lower bound, b is the upper bound, c is the lowest pixel intensity in a range of a-b, and d is the highest pixel intensity in a range of a-b.

4.3 Segmentation

The segmentation process was performed with texture-based segmentation and first-order local entropy [26]. The entropy method can detect the location of cancer tissue by distinguishing objects from the background. In this stage, Ω_k was defined as a small region with window width ($M_k \times N_k$) in the input image, which can be defined in Equations 2 and 3,

$$E\left(\Omega_k\right) = -\sum_{i=0}^{L-1} P_i \log\left(P_i\right) \tag{2}$$

$$P_i = \frac{n_i}{M_k - N_k} \tag{3}$$

where i is the grayscale value, P_i is the grayscale probability of i that appears in the k region, n_i is the number of grayscale pixels i, L is the maximum grayscale, and $E(\Omega_k)$ is the local entropy. This computation used a square-shaped structuring element of 9x9.

The average value of an entropy image was used to produce a binary image with thresholding. In this binary image, the normal tissue became the foreground while the object is the background. Therefore, the inverting method must be performed so that the object becomes the foreground. This was applied by changing the intensity value from 1 to 0, and vice versa. After that, morphological operators such as morphological opening and morphological closing were applied with the oval shape of structuring elements. Morphological closing was used to remove labels on the background and morphological opening was used to remove small objects.

4.4 Feature Extraction

In this research, the border following algorithm [27] was used as feature extraction. This algorithm was applied for obtaining the contour of detected cancer tissue. The output was coordinate points that surround the detected object. Based on this result, the extreme points and objects' moments were calculated. The object's extreme points were calculated for ROI and radius measurement. Extreme points were obtained by finding the outermost coordinate points from the right, left, top, and bottom of the object. Meanwhile, the object's moment was obtained by the Equation 4-6 [28],

$$\mu_{ij} = \sum_x \sum_y (x - x')^i (y - y')^j a_{xy} \tag{4}$$

$$x' = \frac{m_{10}}{m_{00}}, y' = \frac{m_{01}}{m_{00}} \tag{5}$$

$$m_{00} = \sum_x \sum_y a_{xy}, m_{10} = \sum_x \sum_y x a_{xy}, m_{01} = \sum_x \sum_y y a_{xy} \tag{6}$$

where μ_{ij} is the center moment, x is the length of the image, y is the width of image, a_{xy} is the intensity value in points x and y, and (x',y') is the center of coordinate.

Those coordinates were used to obtain ROI coordinate of detected cancer tissue. Then, one of these ROI coordinates was used to obtain the radius of the detected cancer tissue along with the object's moment by measuring the distance between two points as shown by Equation 7.

$$Radius = \sqrt{(x_1 - x_2) + (y_1 - y_2)} \tag{7}$$

4.5 Performance Evaluation

The performance of the research was evaluated using two methods. The first method was to determine the evaluation of breast cancer detection The second method was to determine the evaluation of radius and pectoral muscle measurements. The results evaluation of breast cancer detection was used confusion matrix, which consists of accuracy, precision, sensitivity or recall, and specificity using Equations 8-11,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$\Pr ecision = \frac{TP}{TP + FP} \tag{9}$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (10)$$

$$Specificity = \frac{TN}{TN + FP} \quad (11)$$

where TP, TN, FP, and FN are True Positive, True Negative, False Positive, and False Negative detections, respectively.

The evaluation to find out if there was an increase in accuracy using the MLO view framework was done. The system was tested on an MLO view image without the pectoral muscle and with pectoral muscle.

5. EXPERIMENT AND ANALYSIS

5.1 Pre-processing

In the MLO view image, image orientation and pectoral muscle removal were performed. In the process of removing the pectoral muscle, contour detection was done that consist of Gaussian filter, canny edge, Sobel filter, and Hough lines transform to create a pectoral muscle line. According to [18], the optimal angular limit of the pectoral muscles in mammographic images was 10° to 80°. However, the angular limit in this research was obtained from 10° to 70°. Furthermore, all of the detected lines are in the range of 10° to 70°, the line closest to the edge was chosen and then 160 pixels were added for the x-axis intersection and 110 pixels for the y-axis intersection. This line was defined as the pectoral muscle line to be removed.

Table 1. MSE and PSNR values using different clip limits of CLAHE

Dataset	Clip limit	MSE	PSNR (dB)
CC view: B_3065_1.RIGHT_CC	2.0	63,78	30,08
	6.0	119,82	27,78
	12.0	108,36	27,34
MLO view: A_1530_1.RIGHT_MLO	2.0	34,80	32,71
	6.0	60,75	30,29
	12.0	142,05	26,61

The MLO view without the pectoral muscles and CC view were processed using CLAHE to improve the quality of the image. It was determined to use the structural element size set to 8x8 and the clip limit value of CLAHE was adjusted to 2.0. Comparisons between MSE and PSNR values of raw images using different clip limit is shown in Table 1. Based on Table 1, we concluded that the greater clip limit value used affected to increase the contrast of the image. Clip limits of 6.0 and 12.0 had PSNR values below 30 dB. The results indicated that the output image had low similarity to the original one. The distortion of the image was high. In addition, the MSE values at clip limits of 2.0, 6.0, and 12.0 increased respectively.

The value of the upper bound and the lower bound was critical to the output image in contrast-stretching. The greater the entropy value of a histogram, the intensity of the image was getting more diverse. Furthermore, it also corresponded to the standard deviation value. The high distribution of gray level supported the segmentation method with local entropy because this method used probability value in the small region (Ω_k).

5.2 Segmentation

There are three steps in the segmentation process: calculating local entropy, thresholding, determining size of morphological opening and closing. Entropy was a measurement of both the complexity of the image and the quantity of information that was included in the distribution data. In this study, first-order local entropy was used. There was a correlation between the size of the structuring element and the detection outcome. Entropy with a structuring element of 9x9 was used to obtain optimal cancer detection results. The greater the entropy number, the blurrier the image seemed as the structure's size increased.

Determination of the threshold was done with information from the average value and the maximum value of the entropy results. In the previous study [5] the optimum threshold had not explained, so the experiment on each image used a different threshold value. In this study, the selected threshold value was 3.3 for the 9x9 structuring element size for all CC and MLO views. Detection of suspected cancer was significantly improved by this threshold and kernel size. The higher the threshold value, the more cancer tissue is found, and vice versa. After thresholding was applied, the output image became a binary image with a background value of 0, breast tissue value of 1, and abnormal tissue value of 0. Therefore, it was necessary to perform an inverting process to foreground abnormal breast tissue.

The morphological operator was the next step. Morphological closing and morphological opening were the two kinds of morphological operators employed in this research. Morphological closing was used to remove labels on mammography images, while the morphological opening was applied to remove small objects detected as noise. This method was performed to avoid cancer detection errors and ROI detection errors. The most optimum morphological closing used an elliptical structuring element with a size of 13x13. After morphological closing, the CC and MLO view images were implemented morphological opening using an elliptical structuring element with a size of 9x9. The comparison of the results using morphological closing and opening can be seen in Figure 3. Although the difference in the results of each morphological operator is not clearly visible, this stage greatly affects the next results.

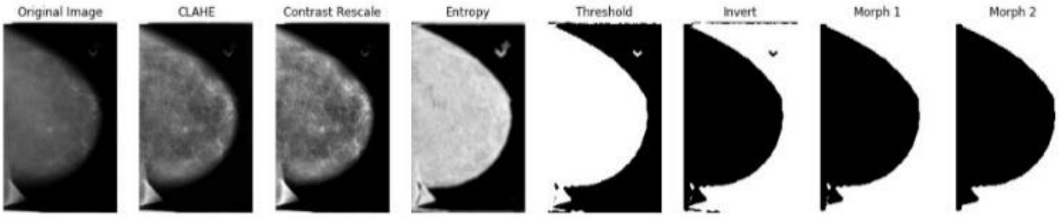

Figure 3. The stages of pre-processing and segmentation in B_3621_1.LEFT_CC.

5.3 Feature Extraction

In this research, the border following algorithm was used as feature extraction. This algorithm was applied for obtaining the contour of detected cancer tissue. The outcome of this method was a set of ROI coordinates that encircle the item suspected of having cancer. After that, the search for extreme points and center moments on the object was calculated. Extreme points were obtained by finding the outermost coordinate points from the right, left, top, and bottom of the object. In the meanwhile, a single point was employed to determine ROI's central moment. Based on Equation 7, one of the points obtained at the extreme point with the center moment resulted in a distance, where this distance was used as the radius value. The findings of the radius calculation might be used to estimate the area of the item suspected of having cancer.

Figure 4(a) presents the findings obtained from the examination of the CC view picture for the identification of breast cancer. Center moment in the CC view dataset (B_3065_1.RIGHT_CC) was obtained in the coordinate of (305, 355) as shown in Figure 4(b). According to these points, the distance was 48 pixels or 2.4 centimeters. Furthermore, the calculated area of the object inside the contour was 40580 pixels. The results of breast cancer detection in the MLO view dataset are shown in Figure 4(c). The MLO view image with pectoral muscle removal can accurately detect the suspected cancer area compared to the result without removing the pectoral muscle as shown in Figure 4(d). Ground truth from dataset B_3065_1.RIGHT_CC and C_0011_1.RIGHT_MLO can be seen in Figures 4(e) and 4(f) respectively.

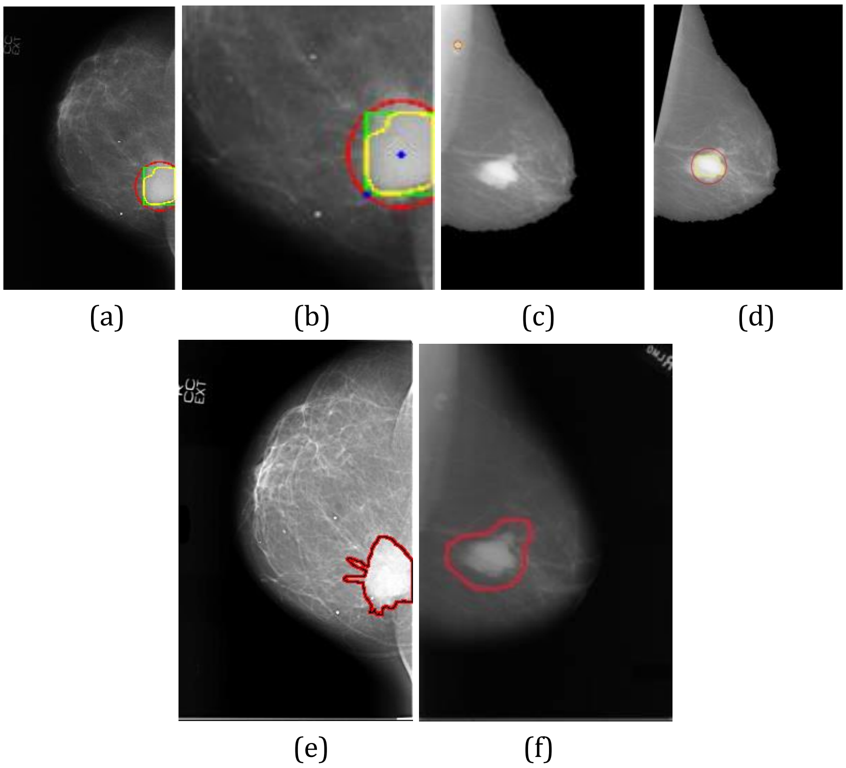

Figure 4. (a) The results of contour detection (yellow), ROI (green), and object circle (red) in dataset B_3065_1.RIGHT_CC, (b) Calculation of cancer radius in object moment and ROI, (c) Breast cancer detection results on MLO view without pectoral muscle in dataset C_0011_1.RIGHT_MLO, (d) Detection results on MLO view with pectoral muscle, (e) Ground truth from dataset B_3065_1.RIGHT_CC, and (f) Ground truth from dataset C_0011_1.RIGHT_MLO.

Performance evaluation for breast cancer detection had been carried out in 120 CC view images and 60 MLO view images as shown in Table 2. Based on the results, this study had sensitivity exceeding the average of 80% of the previous study [5][8][9]. The results of precision, accuracy, and specificity in the MLO view images using pectoral muscle removal have a higher value than the MLO view images with pectoral muscles as shown in Table 2. This is due to the similarity in density between the pectoral muscle and the probable cancerous tissue. For this reason, it is important to segment the pectoral muscle before processing any images. However, there were images in which the segmentation of the pectoral muscles was imperfect. Consequently, they may still lead to mistakes in cancer detection. False positives occurred when the image had high contrast, so in the image enhancement process, there was an overlap between cancer tissue and normal tissue. Normal tissue in the image was detected as cancer tissue. Another cause of false-positive was the pectoral muscle could not be completely segmented. Meanwhile, false negatives occurred because the image had too low contrast so that the suspected cancer tissue could not be detected.

Table 2. Performance evaluation of breast cancer detection

Performance	CC views	MLO views with pectoral muscle	MLO views using pectoral muscle removal
Accuracy	92.0 %	33.6%	68.7%
Precision	97.0%	30.0%	64.7%
Sensitivity	88.0%	80.4%	80.5%
Specificity	96.0%	11.5%	57.1%

In CC view images, there were two images with false-positive results. The proposed structuring element size was very big for covering the cancer area. This caused entropy had a high value then the threshold could not detect cancer tissue area. False negatives in CC view images occurred because the fibroglandular tissue appeared more contrast to other breast tissue. As a result, the proposed algorithm detected fibroglandular tissue as cancerous tissue. In addition, normal tissue with a low intensity of gray level within the structuring element area could be considered as cancerous tissue because it had low entropy value.

6. CONCLUSION

This study aimed to detect breast cancer, improve performance, and get radius and ROI from suspected breast cancer using CC view and MLO view of mammographic images. The study establishes that the accuracy of the proposed framework outperformed other methods previously. The framework in the pre-processing stage can improve the local entropy-based segmentation performance. Local entropy-based segmentation was proven to be able to differentiate the cancer tissue and normal tissue. The entropy threshold value has a significant effect on the detection result. The greater value of the threshold, the more area was detected, and vice versa. Border following method was proven to obtain contours and detect ROI. This approach has also been shown to evaluate the radius and region of probable malignancy. The removal of the pectoral muscle from the MLO view picture has a significant impact on the outcomes of breast cancer diagnosis. Based on the research, the framework was able to detect breast cancer with a sensitivity of more than 80%, both in CC view and MLO view images. It is recommended that further research be undertaken in improving the detection of breast cancer area. Furthermore, the framework can be developed using a convolutional neural network for cancer classification and pectoral muscle segmentation to increase performance.

Acknowledgments

The author(s) disclosed receipt of the following financial support for the publication of this article. This work was supported by PUI-PT AIHeS (Artificial Intelligence for Healthcare and Society), Institut Teknologi Sepuluh Nopember, ITS.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, **Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries**, *CA Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[2] W. A. Berg and J. W. T. Leung, **Diagnostic Imaging Breast**, *Elsevier*, Ed. 3, pp. 4-6, 2019.

[3] L. W. Bassett, M. C. Mahoney, S.K. Apple, and C. J. D'Orsi, **Breast Imaging**, *Saunders Elsevier (Philadelphia)*, Ed. 1, pp. 25-45, 2011.

[4] A. B. R. Arafah and H. B. Notobroto, **Factors Related to the Behavior of Housewives Performing Breast Self-Examination**, *The Indonesian Journal of Public Health*, vol. 12, pp. 143-153, 2017.

[5] A. S. A. B. Sama and S. M. S. Baneamoon, **Breast Cancer Classification Enhancement Based on Entropy Method**, *International Journal of Engineering and Applies Computer Science*, vol. 2, pp. 267-271, 2017.

[6] G. B. Junior, S. V. Rocha, J. D. Almeida, A. C. Paiva, A. C. Silva, and M. Gattass, **Breast Cancer Detection in Mammography Using Spatial Diversity, Geostatistics, and Concave Geometry**, *Multimedia Tools and Applications*, vol. 78, pp. 13005-13031, 2019.

[7] B. Ozlu, M. Sahin, K. E. Erdogan, and B. Kocer, **The Role of Radiography Guided Non-Palpable Breast Lesion Marking On the Diagnosis of Early Breast Cancer and Proliferative Diseases**, *Van Tip Derg Medical Journal*, vol. 25, no.1 pp. 34-42, 2018.

[8] N. M. Basheer and M. H. Mohammed, **Segmentation of Breast Masses in Digital Mammograms Using Adaptive Median Filtering and Texture Analysis**, *International Journal of Recent Technology and Engineering*, vol. 2, no. 1, pp. 39-43, 2013.

[9] A. K. Singh and B. Gupta, **A Novel Approach for Breast Cancer Detection and Segmentation in a Mammogram**, *Procedia Computer Science*, vol. 54, pp. 676-682, 2015.

[10] A. P. Singh and B. Singh, **Texture Features Extraction in Mammograms Using Non-Shannon Entropies**, *Machine Learning and Systems Engineering. Lecture Notes in Electrical Engineering*, Dordrecht, Springer, pp. 341-351, 2010.

[11] E. A. Sickles, C. J. D'Orsi, and L. W. Bassett, **ACR BI-RADS Atlas, Breast Imaging Reporting and Data Systems**, *Reston: American College Radiology*, 2013.

[12] S. Sasikala, M. Ezhilarasi, and S. A. Kumar, **Detection of Breast Cancer Using Fusion of MLO and CC View Features Through a Hybrid Technique Based on Binary Firefly Algorithm and Optimum-Path Forest Classifier**, *Applied Nature-Inspired Computing: Algorithms and Case Studies, Springer Tracts in Nature-Inspired Computing (STNIC)*, 2020.

[13] J. R. G. Sepulveda, M. T. Cisneros, D. A. M. Arrioja, J. R. Pinales, O. G. I. Manzano, G. A. Cervantes, and A. G. Parada, **Digital Image Processing Technique for Breast Cancer Detection**, *International Journal Thermophys,* vol. 34, pp. 1519-1531, 2013.

[14] A. Unni, N. Eg, S. Vinod, and L. S. Nair, **Tumour detection in double threshold segmented mammograms using optimized GLCM features fed SVM**, *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, IEEE, pp. 554-559, 2018.

[15] S. Sasikala and M. Ezhilarasi, **Fusion of k-Gabor Features from Medio-Lateral-Oblique and Craniocaudal View Mammograms for Improves Breast Cancer Diagnosis**, *Journal of Cancer Research and Therapeutics*, vol. 14, pp. 1036-1041, 2018.

[16] S. A. Taghanaki, Y. Liu, B. Miles, and H. Ghassan, **Geometry Based Pectoral Muscle Segmentation from MLO Mammograms Views**, *IEEE Transactions on Biomedical Engineering*, vol. 64, pp. 1-11, 2017.

[17] R. Ferrari, R. Rangayyan, J. Desautels, R. A. Borges, and A. Fere, **Automatic Identification of the Pectoral Muscle in Mammograms**, *IEEE Transaction on Medical Imaging*, vol. 23, pp. 232-245, 2004.

[18] P. Shi, J. Zhong, A. Rampun and H. Wang, **A hierarchical pipeline for breast boundary segmentation and calcification detection in mammograms**, *Computers in Biology Medicine*, vol. 96, pp. 178-188, 2018.

[19] A. R. Beeravolu, S. Azam, M. Jonkman, B. Shanmugam, K. Kannoorpatti, and A. Anwar, **Preprocessing of Breast Cancer Images to Create Datasets for Deep-CNN**, *IEEE Access*, vol. 9, pp. 33438-33463, 2021.

[20] M. A. Al-masni, M. A. Al-antari, J. M. Park, G. Gi, T. Y. Kim, and P. Rivera, **Detection and Classification of the Breast Abnormalities in Digital Mammograms via Regional Convolutional Neural Network**, *The 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Seogwipo, 2017.

[21] K. Akila, L. S. Jayashree, and A. Vasuki, **Mammographic Image Enhancement Using Indirect Contrast Enhancement Techniques**, *Procedia Computer Science*, vol. 47, p. 255 – 261, 2015.

[22] S. Chen and A. R. Ramli, **Contrast enhancement using recursive mean-separate histogram equalization for scalable brightness preservation**, *IEEE Transactions on Consumer Electronics,* vol. 49, no. 4, pp. 1301-1309, 2003.

[23] P. Carneiro, C. Debs, A. Andrade, and A. Patrocinio, **CLAHE Parameters Effects on the Quantitative and Visual Assessment of Dense Breast Mammograms**, *IEEE Latin America Transactions*, vol. 17, no. 5, pp. 851-857, 2019.

[24] **Curated Breast Imaging Subset of DDSM**, https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM, Accessed on 05/01/2021.

[25] A. A. Kayode, B. S. Afolabi, and B. O. Ibitoye, **An Explorative Survey of Image Enhancement Techniques Used in Mammography**, *International Journal of Computer Science Issues*, vol. 12, no. 1, pp. 72-79, 2015.

[26] M. Pawar and S. Talbar, **Local entropy maximization based image fusion for contrast enhancement of mammogram**, *Journal of King Saud University – Computer and Information Sciences*, vol. 33, pp. 150-160, 2021.

[27] K. Wu, W. Li, Y. Ge, Y. Chen, Y. Shen, J. Hu, L. Wang, K. Chu, B. Liu, J. Yan, **A subpixel edge detection algorithm based on the combination of border following and gray moment**, IEEE 10th International Conference on Nano/Molecular Medicine and Engineering (NANOMED), pp. 34-38, 2016.

[28] J. T. D Bertsimas, **Introduction to linear optimation**, *Massachusetts: Athrns Scientific*, 1997.

[29] T. M. Fahrudin, I. Syarif, and A. R. Barakbah, **Data Mining Approach for Breast Cancer Patient Recovery**, *EMITTER International Journal of Engineering Technology*, vol. 5, no. 1, pp. 36-71, 2017.

[30] N. P. Firdi, T. A. Sardjono, and N. F. Hikmah, **Using Pectoral Muscle Removers in Mammographic Image Process to Improve Accuracy in Breast Cancer**, *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, vol. 55, pp. 131-142, 2022.