Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CCC: AN APPROACH FOR DETECTION THE MASS OF BREAST CANCER IN MAMMOGRAM IMAGES

[1] MOHAMMAD ALFRAHEED

[1]Department of Computer Science and Information Technology, Faculty of Science, Tafila Technical University, Jordan

E-mail: [1]alfraheed@ttu.edu.jo

ABSTRACT

Due to the need to early detection of the breast cancer, the mammogram images are taken for cancer patient. The poor visibility and weak contrast are usually highlighted as challenges in such mammogram images. These challenges decrease normally the accuracy in computerized mammogram image segmentation. In this work, an effective method has been developed to detect and extract the cancer mass for breast cancer from mammogram images. The proposed method has been developed to first, extracting the breast area (i.e. foreground area) from other objects in mammogram image (i.e. background area). Second, the Mahalanobis distance value has been introduced as a solution for detecting the cancer mass from objects in breast area. The idea behind using the Mahalanobis distance value is the need to decrease the computation complexity and increase the detection accuracy. In addition, the running time of the proposed method has to be reduced. Therefore, the proposed method reduces the running time to approximately 1.2 second. Compared to other methods, the proposed method reduces the running time and avoids the training stage of the cancer mass detection. Furthermore, the proposed method does not resize the original mammogram image in order to keep the original details of the cancer mass.

Keywords: *Breast Cancer, Mammograms Images, Segmentation, Detection*

1. INTRODUCTION

This guide provides details to assist authors in preparing a paper for publication in JATIT so that there is a consistency among papers. These instructions give guidance on layout, style, illustrations and references and serve as a model for authors to emulate. Please follow these specifications closely as papers which do not meet the standards laid down, will not be published.

Cancer masses cause usually uncontrolled multiplication in any area of the human body. Some people whom we love die due to the cancer disease. Breast cancer is the second leading causes of the cancer related death among women [1]. In Jordan, the ministry of health encourages always all of women to diagnose the disease because of increasing the probability of surviving the patient [2]. Women who are more than 60 years old could be diagnosed as patient of the breast cancer. 37% of breast cancer cases are detected for patient whose age is more than 65. In contrast, only 14% of cases are detected under the age of 45 [3].

The digital images generated through the process of the mammography are called Mammograms. The mammogram is usually the first tools to diagnose patients; especially those suffer from the breast cancer. Although the digital mammograms images plays an important role in early detection of the cancer mass, the poor visibility encourages researchers to develop techniques for enhancing the visibility and mammograms images segmentation [4].

Generally, the image segmentation is a technique to extract automatically, the region of interest from the input image [5]. Computerized mammograms images segmentation is a permanent challenge due to poor resolution and weak contrast of mammograms images[6]. Several of algorithms had been developed for extracting the cancer mass from the digital mammograms images. These algorithms and techniques vary in term of advantages and drawbacks.

In this work, an effective method has been developed to detect the cancer mass of the breast cancer in the mammograms images. The proposed method has to be applicable for different types of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

mammogram images (i.e. MedioLateral Oblique and CranioCaudal) . Furthermore, the proposed method has to overcome the high computational complexity.

The rest of the paper is organized into five sections as follows; the section 1 gives the introduction. Section 2 discusses the recent researches in mammogram images detection. The section 3 elaborates the proposed method. Section 4 shows the results and discussion for the proposed methods. Finally, the last section includes the conclusion and future work (section 5).

2. RELATED WORK

The recent researches for the detection of the breast cancer are being developed towards joining detection and classification in term of the neural network algorithms [1] [6] [7]. Richard platanio et al [8] have proposed a framework for automated breast cancer detection and diagnosis using convolutional neural networks. The latter has been also used in [9] to develop a CNN-based method for automatic mass detection and classification in mammograms. Moreover, the Multi-Support vector machine (SVM) has been used in [10] not only to detect and classify the mass abnormalities but also to increase the diagnostic accuracy of the optimum classification . The generated results are applicable in the Computer Aided Diagnosis (CAD) systems. However, the false-positive results are usually generated due to the low accuracy of the detection process and the quality of the extracted features.

The current researchers pay close attention for the pre-processing and initial segmentation steps. The idea behind this attention is to have low computational complexity for different kinds of medical images and to have high quality segmentation. In [11], a parameter-adaptive pulse-coupled neural network had been generated using an optimal histogram threshold. In contrast, Yang Jiao et al [12] have generated sub-regions and their features based on watershed transform algorithm in order to be used in a network clustering model.

In this work, a method has been developed to extract automatically, the region of cancer mass and to generate precisely the required features and parameters. The latter are used in advance by one of the neural network algorithm to classify the cancer mass.

In image segmentation, the most important contribution is how to extract precisely the object of interest regardless its shape, position and appearance. Moreover, medical images segmentation techniques are being developed to provide accurate information about object of interest (i.e. size, location, shape characteristics … etc.). Recently, a combined approach has been developed for mammograms images segmentation based on watershed transform and k-means clustering [13]. Since the watershed transform techniques are usually affected by local intensity variation and they are controlled by initial position about object of interest, hang min et al [14] have developed a novel approach for detecting and segmentation breast masses based on multi-scale morphological filtering and a self-adaptive cascade of the random forests. The detection based multi-scale morphological had been developed to generate candidate regions. Therefore, Kanchan et al [15] have presented an approach for mass detection in mammograms which is based on the variational level set function. However, it is computationally expensive. This kind of computation has been arisen due to low contrast and complicated structured background. To overcome this challenge, a method has been developed in [16] to enhance the global and local contrast of the given image in the first hand. On the other hand, the segmentation breast regions had been processed to obtain a better visual interpretation, analysis and classification of mammogram masses. In contrast, Shiju Yan et al [17] have been developed a method in which three types of image features have been calculated (i.e. a symmetry, mean and the maximum of the images features). These features have been fused in two stage classification schema to generate a bilateral mutual threshold. The latter, in turn, improves the detection accuracy of the breast cancer region. Using threshold or multi-threshold in segmentation is a novel technique. In [18], a method has been developed to establish specific threshold values for background and object of interest during the initialization stage. Although the detection based threshold values plays an important role in the detection process, the false-positive pixels have to be processed in advance to emphasis the high accuracy of detection. This kind of false positive pixels generates usually a sub-group corresponding to the less representative pixels of object. Therefore, Ojeda-Magana et al [19] have developed a method to improve the detection of such pixels and to recognize clearly between pixels related to object of interes

t and non-related pixels. Their developed method depends on the Possibilistic Fuzzy c-Means

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(PFCM) clustering is computationally expensive and it needs a training stage for accurate clustering.

In this Work, a method has been proposed to detect automatically the cancer mass without training stage and initial parameters. Furthermore, the proposed method has to take in its account the high computational complexity of such kind of method. Meaning that, the running time of the proposed method has to be short as much as it can. Normally, the size of the mammogram images is so large because the image shows very small details of the patient. Therefore, the size of the mammogram image has to be processed here as it is given by the original size without resizing process.

3. PROPOSED METHOD

The proposed method consists of four stage; extracting the breast region, local contrast enhancement for breast region, breast region segmentation and detection of the cancer mass. Figure 1 shows the flow diagram of the proposed method. Each stage is described in detail below.

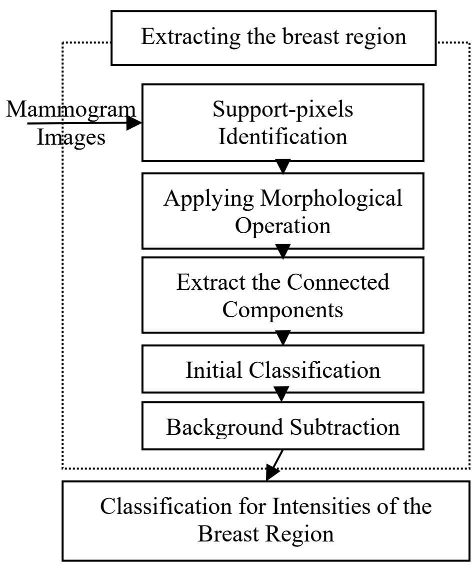


Figure 1. Flow Diagram of the Proposed Method

3.1. Extracting the breast region

Beside of the breast region (i.e. Foreground Area), the mammogram image shows other information (i.e. labels, patient's information and noises). It is usually called background Area. Gray values associated with these information affects negatively on the detection results. Therefore, the breast region has to be precisely extracted for processing in next stages. Furthermore, the size of the mammogram images is normally high due to mechanism of the x-ray scanning. The memory used for implementation is consequently increased. Processing only the foreground reduces consequently the memory consumption and the computation cost. A proposed approach has been introduced to extract the breast region. The steps of the proposed approach are described as follows;

3.1.1. **Support-*pixels Identification***

Here, the boundaries of image's object are determined. Determining these boundaries enables the proposed method to generate separated objects for foreground and background area. Since the canny detector of the edge detection has been effectively used in medical image boundaries detection [18], it has been introduced here as a preferred detector. The canny detector is calculated by the following equations [5].

$$G = \sqrt{{G_x}^2 + {G_y}^2} \qquad \ldots (1)$$

$$\theta = \arctan\left(\frac{G_x}{G_y}\right) \qquad \ldots (2)$$

$$G_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} \qquad \ldots (3)$$

$$G_y = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} \qquad \ldots (4)$$

Where:

G: The gradient strength

θ: The direction of the gradient strength

G_x: The gradient in x-direction

G_y: The gradient in y-direction

Figure 2-A shows a sample of the mammogram images. The 3-dimensional gray value distribution is shown in Figure 2-B for foreground and background area. Figure 2-C shows the 3-dimensional gray values for only foreground intensities. As a consequence of the canny detector, the Figure 2-D shows the detected boundaries for all objects in mammogram image.

Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

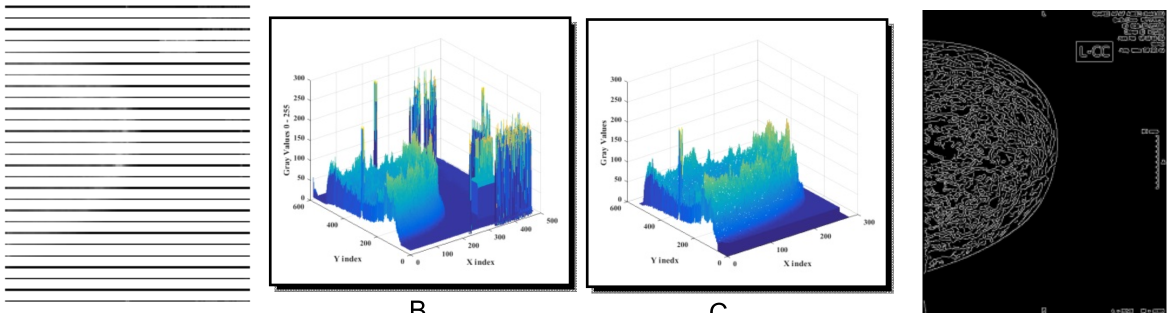


Figure 2. Different views of the mammogram images. (A) The original view of the mammogram image. (B) 3-dimensional gray level distribution of the foreground and background area. (C) 3-dimensional gray level distribution of the foreground area. (D) The detection boundaries of the whole image based on canny detector

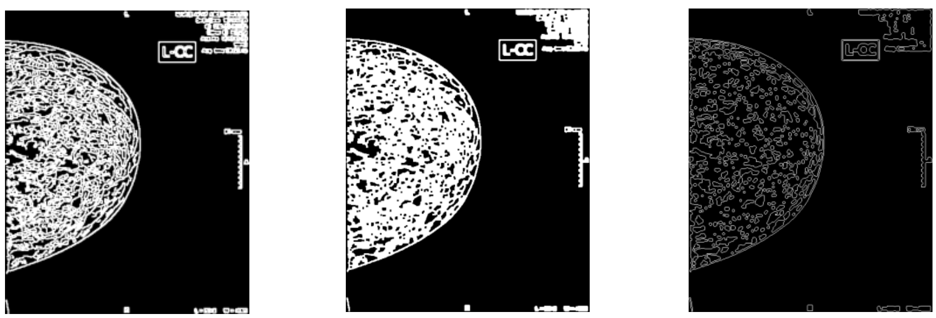


Figure 3. Outputs generated from the Morphological Operations. (A) Outputs generated from the dilation operation. (B) Outputs generated from the closing operation. (C) Outputs generated from the removing operation

3.1.1. Applying Morphological Operation

Some pixels generated by canny detector are not considered to be a part of the object boundary. These pixels are removed by applying morphological operations. Three morphological operation operations have been applied for image of the object boundaries, which are Dilation, Closing and Removing [5]. Figure 3- A, Figure 3- B and Figure 3- C show the results generated by dilation, closing and removing respectively.

3.1.2. Extract the Connected Components

After having the black-white mammogram images (i.e. Binary Image). It is necessary to extract the connect object of interest. This extraction enabled the proposed method to emphasize its processing on those objects instead of discarded and useless objects. Defining connectivity in binary image classifies its pixels as 1 when the pixel is a part of pattern. In contrast, pixels associated with 0 are classified as a part of background. Generally, two types of connectivity are introduced 4-connectivety and 8-connectivety [5]. The latter has been used here as a default parameter for connectivity. Figure 4 shows the results generated by extracting all of the connected objects.

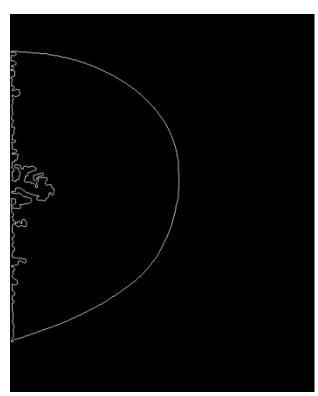


Figure 4. The maximum connected object extracted from the mammogram image

3.1.3. Initial Classification

Within this step, the foreground and background are automatically classified based on the two factor;

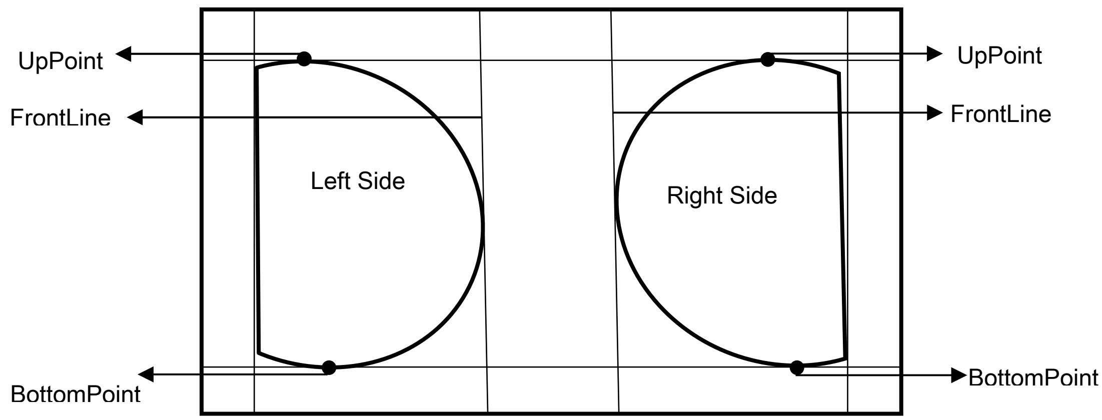


Figure 5. The flow chart for variables used in the initial classification

The connected objects (IM_{CONN}) and support-pixels (IM_{SP}) identified by the previous steps. Both factors play an important role in initial classification because they identified the foreground region (breast area) and background region. Beside, pixels associated with the breast region are more meaningful than other pixels in the mammogram image. The Figure 5 shows the position of the variables used here.

Furthermore, the mammogram images are captured for both side of breast (i.e. Left and Right side). The step of the initial classification has to be able to classify the position of the breast side (i.e. Left or Right side). The initial classification has been implemented based on the following equations.

$$A_{n\times m} = \begin{cases} 1....\forall IM_{SP} \&\& \forall IM_{CONN} \\ 0.....O.W \end{cases} \quad \dots (5)$$

In equation 5, A_{n*m} is a binary matrix with n height and m width. Both images IM_{SP} and IM_{CONN} are merged to matrix A. If the pixel in IM_{SP} or in IM_{CONN} has been marked as a pixel of interest the corresponding cell of A is filled out by 1. Otherwise the cell of A is filled out by 0.

$ROWs = \{i, \quad iff\ A[i][:] = 1\}$ … (6)

In equation 6, ROW_S is a vector of Size S. The vector contains the indexes of rows for matrix A. The cell of vector is filled out by the index of row if the row of matrix A contains 1. Otherwise the cell is filled out by an ignored value.

$COLz = \{j, \quad iff\ A[i][:] = 1\}$ … (7)

In equation 7, COL_z is a vector of Size Z. The vector contains the indexes of columns for matrix A. The cell of vector is filled out by the index of col if the col of matrix A contains 1. Otherwise the cell is filled out by an ignored value.

[UpPoint IndexMin] =
min(Row[index]), …. $\forall$ index …(8)

After generating the vector Row whose elements have been addressed from 1 to index the row's index of UpPoint has been calculated using the equation 8. The latter find the minimum row's index in vector Row. The **Error! Reference source not found.** shows the position of the UpPoint.

[BottomPoint IndexMax] =
max(Row[index]), …. $\forall$ index …(9)

In the same way, the BottomPoint has been calculated by equation 9. The BottomPoint has been calculated using the minimum row's index instead of the minimum row's index. The Figure 5 shows the position of the BottomPoint.

BreastPos =
{
LeftSide ,
(Col(IndexMin)ORCol(IndexMax))<(m/2)
RigthSide ,
(Col(IndexMin)ORCol(IndexMax))>=(m/2)
} … (10)

The mammogram images show usually a side of the patient's breast (i.e. Left or Right Side). Here, the mammogram image has to be classified into which

Journal of Theoretical and Applied Information Technology
30th September 2018. Vol.96. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

side it has. Three parameters have been used to initially classify the image, which are the vector Col, IndexMin and IndexMax. In case the column's index for the UpPoint or the BottomPoint is less than the half size of the image width (i.e. m), the patient's breast is the LeftSide. Otherwise, the Patient's breast is the RigthSide.

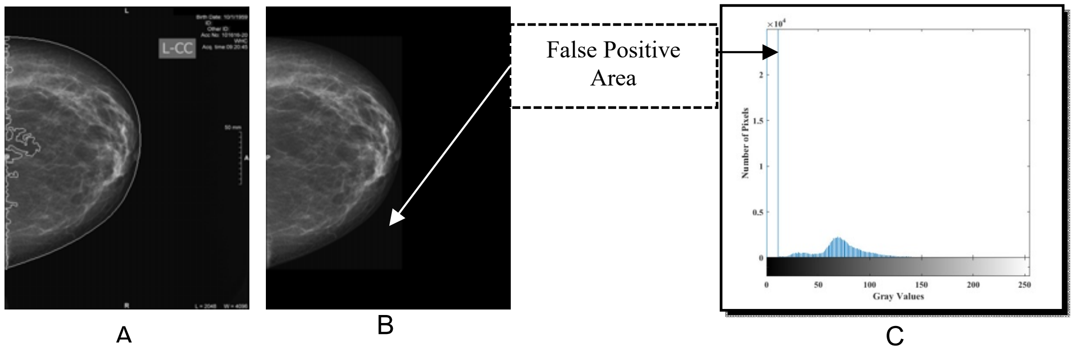


Figure 6. The pixels and their gray values located for background and foreground. (A) The mammogram image with the object of the interest. (B) The object of interest with the false positive area. (C) The Histogram of the gray level of the foreground with the false positive area

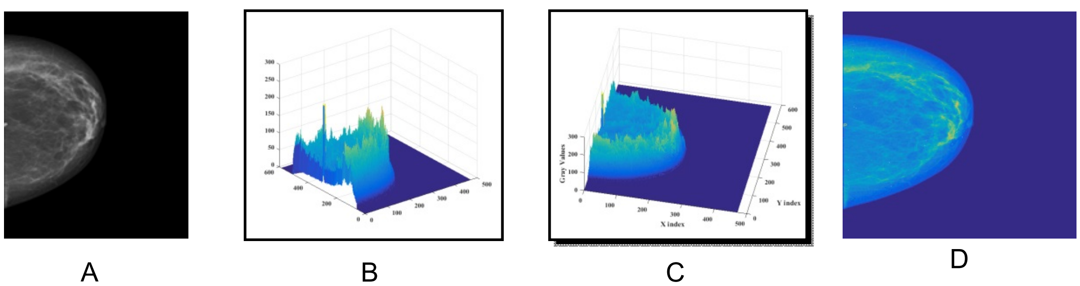


Figure 7. The 3-dimensional views for the object of interest (A) The foreground area in a gray values view (B) The 3-dimensional view based on 45 rotated degrees. (C) The 3-dimensional view based on 15 rotated degrees. (D) The 3-dimensional view based on a vertical point of view

$$\text{FrontLine} = \begin{cases} \min(\text{Col}[index]) & , \forall\ index\ \&\&\ \text{LeftSide} \\ \max(\text{Col}[index]) & , \forall\ index\ \&\&\ \text{RightSide} \end{cases} \quad \ldots (11)$$

The FrontLine (shown in Figure 5) is necessary to be determined because it determines front column of the breast area in the mammogram images. The element of the vector Col has been addressed from 1 to index (i.e. the size of the vector). The FrontLine represents the last column determined in the vector Col. In case the breast is the LeftSide, the minimum column's index is marked as a FrontLine. Otherwise, the maximum column's index is marked as FrontLine.

3.1.4. Background subtraction

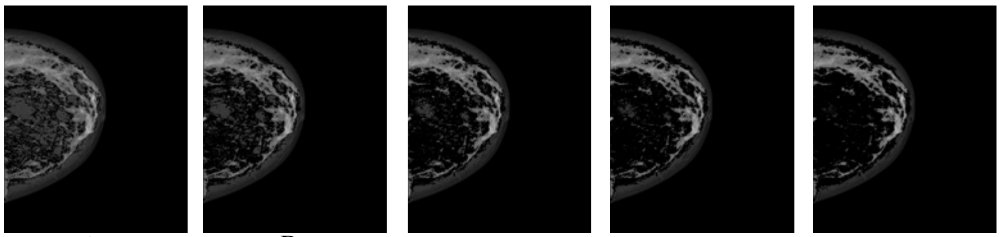
In the last step, pixels are initially classified into background and foreground (i.e. object of interest). The latter is shown in the Figure 6 - A. However, some pixels are marked as a false positive area. These pixels are usually positioned close to object of interest as shown in Figure 6 - B. The background subtraction supports the initial classification in the previous step. The subtraction has been implemented by histogram processing. The latter has been used due to the simplicity of implementation and less computationally expensive [16]. Figure 6 - C shows the histogram of the gray level of the Left side. The false positive area distinguishes itself from other by its number of pixels. It is here determined as the second maximum value of gray level as shown in the Figure 6 - C.

3.2. Classification for Intensities of the Breast Region

As shown in Figure 7 - A, the support-pixels have been only extracted. These support-

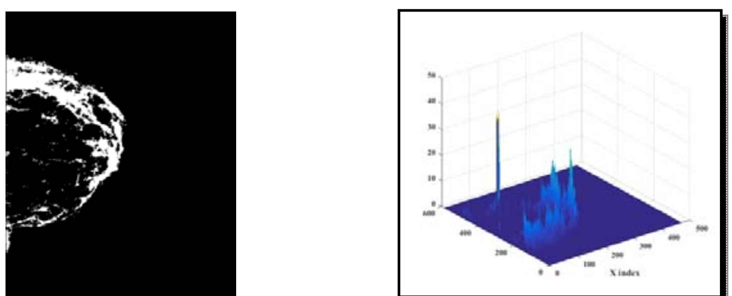
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

pixels have been shown in 3D-Meshgrid. The latter has been rotated to change the point of view as shown in Figure 7 - B, Figure 7 - C and Figure 7 - D. The intensity distribution has been shown as color levels. These levels have to be automatically classified into regions of interest. Moreover, features extraction could be run over each of these regions. To classify these levels, some researches have been applied using the local contrast enhancement [16].



A B C D E

Figure 8. The different gray level views based on the value of the Mahalanobis distance values. (A) The surrounded area assigned to 0. (B) The surrounded area assigned from 0 – 0.1 (C) The surrounded area assigned from 0 – 0.2 (D) The surrounded area assigned from 0 – 0.3 (E) The surrounded area assigned from 0 – 0.4



A B

Figure 9. The extracted cancer mass (A) The extracted area in Black-White view (B) The 3-dimensional view of the cancer area

The drawback of this approach is the patch size of intensity has been decreased as much as possible due to the computational complexity. Here the Mahalanobis distance [20] has been employed for processing and classifying the intensity distribution of the breast region.

Initially, the matrix of support-pixels IM_{SP} is converted to a vector matrix using the following equations;

$$A_{m,n} = IM_{SP} = \begin{Bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{Bmatrix} \quad \ldots(12)$$

$$B = IM_{SP} = \begin{Bmatrix} a_{1,i} & , & 1 \le i \le n \\ a_{2,(i-n)} & , & n+1 \le i \le 2n \\ a_{3,(i-2n)} & , & 2n+1 \le i \le 3n \\ \vdots & \vdots & \vdots \\ a_{(m),i-(m-1)n} & , & (m-1)n+1 \le i \le mn \end{Bmatrix} \quad \ldots (13)$$

The generated vector B has zero's values which represents the black pixels of the discarded pixels. Therefore, these pixels have to be removed from the vector using the following equation.

$$E = B \quad , \text{iff } B[i] \neq 0 \text{ and } 1 \le i \le mn \qquad \ldots (14)$$

The Mahalanobis distance is a statistic measure between a value and a distribution of numerical values [20]. Since the data given here is only a vector of values (i.e. E), the vector is duplicated into two identical groups (i.e. E_1, E_2). Meaning that, the Mahalanobis distance measure the distance between each value in the vector and the distribution of the vector's values. The distance is given based on the following equation;

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$D = (E_1 - M)^T C^{-1} (E_2 - M) \quad \ldots (15)$$

E_1 : The vector of values
M: The vector of mean values of independent variables (E_2)
C^{-1}: Inverse covariance matrix of independent variables (E_2)
In order to extract the area surrounded the cancer area, the Mahalanobis distance has been used as ranges. The surrounded area has been highlighted as a lower range. Whenever the Mahalanobis distance is close to around zero, the surrounded area is highlighted. Therefore, the surrounded area has been assigned to range from 0 to 0.4. Figure 8 shows the surrounded area at different ranges.

Once the surrounded area is extracted, a simple histogram process generates automatically an optimal threshold value for extracted the cancer area as shown in Figure 9.

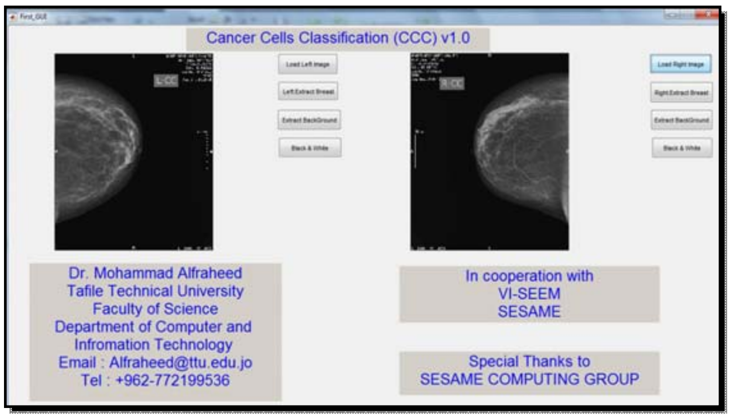


Figure 10. The Main interface used in the EU Project called VI-Seem via CCC

4. RESULTS OF EXPERIMENT AND DISCUSSIONS

5.

Here, the proposed method has been implemented using Matlab. Tools of the image processing box had been fused in the proposed method. The experiment and performance evaluation were carried on the mammograms images taken from the prince zaid bin alhussein military hospital. Images in the database consist of left and right breast image. Some of these images represent a breast cancer case. Other images are for breast without cancer region. Two types of mammogram images has been considered in this study. The first type is mammogram images based on CranioCaudal view (i.e. CC) [21]. The latter is captured when the direction of entry of the x-ray beam enters at the cranial end of the part being examined and exits at the caudal end. In contrast, the second type is mammogram images based on MedioLateral Oblique view (i.e. MLO). The MLO is taken from oblique or angled view. This type is usually preferred than CC because it allows imaging of greatest amount of breast tissue [22].

Figure 10 shows the main interface used to maintain the mammogram image step by step. The execution file of the developed application has been uploaded through project VI-SEEM [23] titled by CCC. The latter has been offered for those interested in medical application of breast cancer.

The proposed method has been successfully run to detect the region of interest (the breast cancer area). Despite of the noise and the intensity distraction in the breast area, the proposed method had extracted the cancer area as shown in Figure 11– E, Figure 11 - F, Figure 11 - G, Figure 11- H. The proposed method has successfully classified the foreground and background for the mammogram images as shown in Figure 11- C, Figure 11 - D. As a consequence of the proposed method, the cancer mass has been extracted in black-white image (i.e. the white area is the cancer mass as shown in Figure 11 - E, Figure 11 - F) and in a gray image as shown in Figure 11- C, Figure 11 - D.

The proposed method has to be adaptable with different view of the mammogram images. *Figure 12* - A and *Figure 12* - B show the mammogram image’s left side and right side respectively in a black-white view. The latter is usually given to extra highlight the cancer area. The proposed method has successfully adapted for the black-white view. The Figure 12 - C shows the successful output of the initial classification for right side of the breast. In contrast, the satisfied results have been shown for the left side of the breast as shown in Figure 12 - B. Although these

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

satisfied results are given by the proposed method, it has successfully detected the cancer area and highlighted it by a gray color for left and right side of the breast as shown in Figure 12 - E, Figure 12 - F.

Some case of the cancer area is not diagnosed as cancer case because the highlighted area is either in early stage or it is generated by other reasons. This case has been also tested by the proposed method. The proposed method has successfully detected the cancer area as shown in Figure 13. However, the proposed method is not able in this study to diagnose in which stage the highlighted area is.

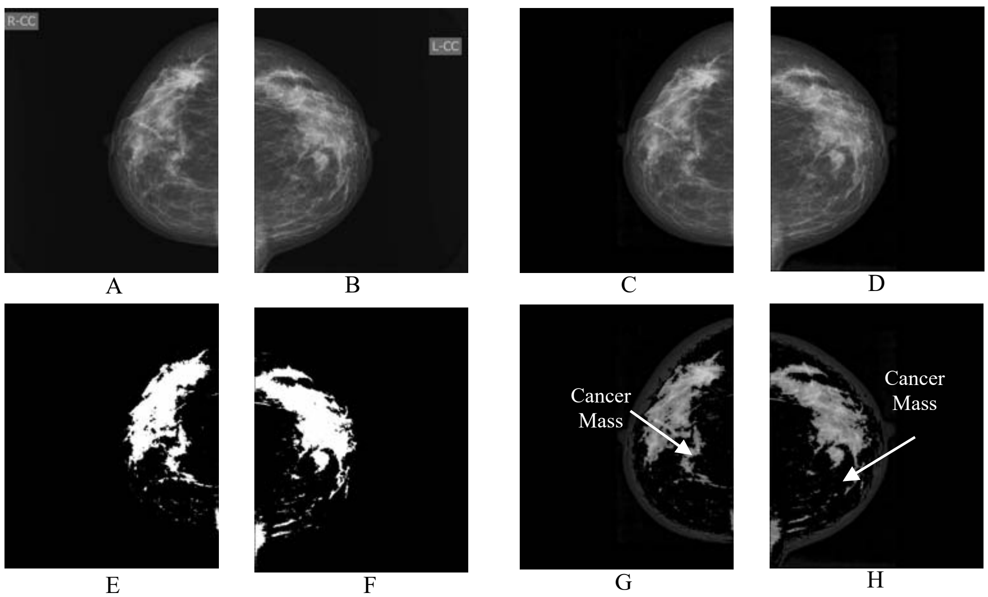


Figure 11. Results generated by proposed method for mammogram image of a negative case based on CC view (A) Right side of breast (B) Left side of breast (C) The foreground area in mammogram image for right side (D) The foreground area in mammogram image for Left side (E) The black-white area of the cancer mass of the right side (F) The black-white area of the cancer mass of the left side (G) The gray value area of the cancer mass of the right side (H) The gray value area of the cancer mass of the left side

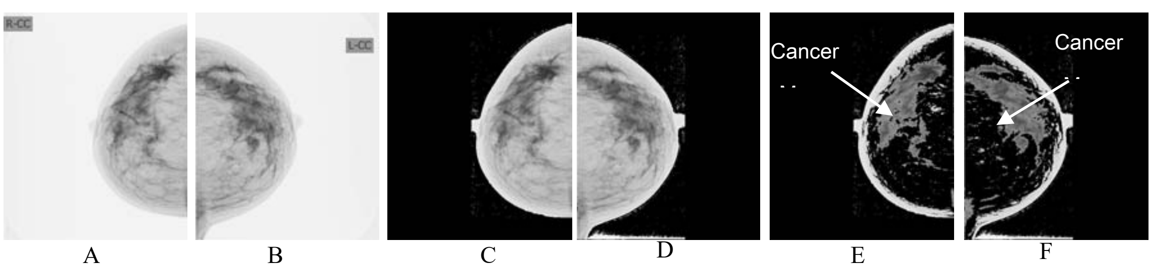


Figure 12. Results generated by proposed method for black-white mammogram image of a negative case based on CC view. (A) Right side of breast (B) Left side of breast (C) The foreground area in mammogram image for right side (D) The foreground area in mammogram image for Left side (E) The gray value area of the cancer mass of the right side (F) The gray value area of the cancer mass of the left side

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

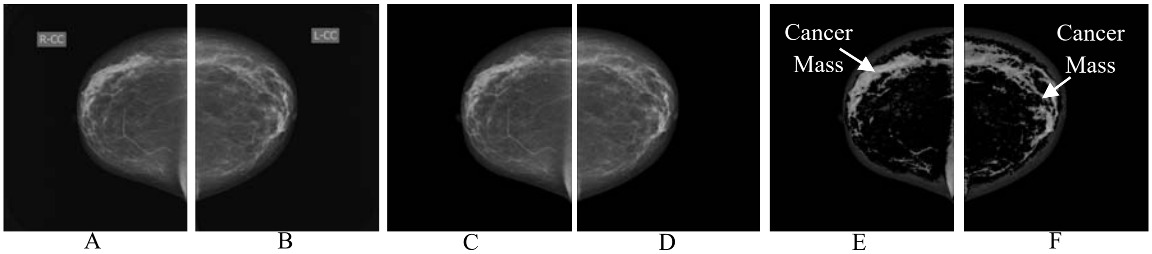


Figure 13. Results generated by proposed method for normal mammogram image of a free case based on CC view. (A) Right side of breast (B) Left side of breast (C) The foreground area in mammogram image for right side (D) The foreground area in mammogram image for Left side (E) The gray value area of the cancer mass of the right side (F) The gray value area of the cancer mass of the left side

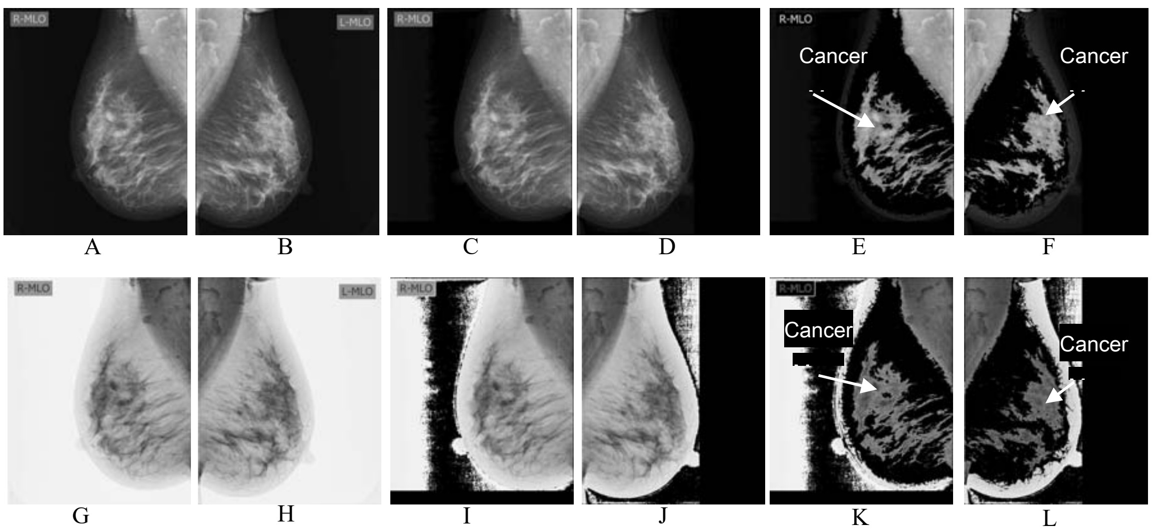


Figure 14. Results generated by proposed method for normal & black-white mammogram image of a negative case based on MLO type. (A) Right side of breast in normal mammogram image (B) Left side of breast in normal mammogram image (C) The foreground area in mammogram image for right side (D) The foreground area in mammogram image for Left side (E) The gray value area of the cancer mass of the right side (F) The gray value area of the cancer mass of the left side. (G) Right side of breast in black-white mammogram image (B) Left side of breast in black-white mammogram image (C) The foreground area in mammogram image for right side (D) The foreground area in mammogram image for Left side (E) The gray value area of the cancer mass of the right side (F) The gray value area of the cancer mass of the left side

Due to the need to show to which extent the proposed method is able to be adapted for different types of mammogram images, the proposed method has been tested over the MLO type. Moreover, the proposed method has been also tested over the black-white image of MLO type. Figure 14 shows these images and the results generated by the proposed method. Figure 14 - E and Figure 14 – F show the detected cancer area of the MLO type which has been highlighted by gray color. In contrast, the detected cancer area based on MLO type and black-white view has been shown in Figure 14 - K and Figure 14 - L.

In order to show the running time of the proposed method, a comparison has been extracted in the Table 1 based on the following parameters: First (case). The mammogram image is for a cancer patient (i.e. Negative) or not (i.e. Free). Second (Side) represents which side has been taken Left breast side (i.e. Left) or Right breast side (i.e. Right). Third (Type) is the type of mammogram image CranioCaudal view (i.e. CC) or MedioLateral Oblique (i.e. MLO). Fourth (Normal) is for a gray value mammogram image. As for the black-white view of mammogram image is given in the last parameter (i.e. Black-White). Running time

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

has been calculated per seconds. As shown in the Table 1, the average of running time is approximately 1.2 second.

Table 1. The running time of the proposed algorithm based on different types, cases and sides for normal and black-white mammogram images

Case	Side	Type	Normal (Sec)	Black-White (sec)
Negative	Left	CC	1.956	1.2279
		MLO	1.3078	1.3626
	Right	CC	1.2348	1.2811
		MLO	1.2900	1.2846
Free	Left	CC	1.2546	1.2649
		MLO	1.8789	1.3602
	Right	CC	1.2153	1.2566
		MLO	1.2661	1.3062

Five parameters are used for the comparison as shown in

Table 2 . The first “Resize” shows whether the method resizes the original mammogram image. The second parameter (“Size”) is the input size of the mammogram image. The third parameter “Training Stage” represents whether the method uses a training classifiers to detect the caner area. The fourth parameter is the “Initial parameter” which shows whether the method needs initial parameters given from the user for running. Finally, “Clustering way” represents the method’s strategy to extract the cancer area from the mammogram images.

Table 2. A Comparison Between The Proposed Method And Other Methods

Method	Resize	Size	Training Stage	Initial parameters	Clustering strategy	Running Time (sec)
Sheng-chih Yang [18]	Yes	1024*1024	Not-required	Yes	PSCSM	-
Kashyap [15]	Yes	1024*1024	Not-required	Yes	BSIF + LBP	≈2.7
Dhungel [7]	Yes	120*120	Required	Yes	Machine Learning Algorithms	-
Kai Hu [24]	Yes	2048*2048	Required	Yes	Machine Learning Algorithms	-
Ojeda-magana [19]	Yes	256*256	Required	Yes	PFCM	-
The proposed method	No	4704*5840	Not-required	No	Mahalanobis distance	≈1.2

Starting from the first parameter, the proposed method does not resize the input image into a small size. Despite the resizing process saves the consuming time and memory, the appearance and the low details features of cancer area are discarded. Instead, the proposed method uses the original size of the mammogram image (i.e. shown in the second parameter). An initial classification process has been applied for automatically classifying the foreground and background of the mammogram images. Therefore, the proposed method is able to decrease the consuming time and memory without changing the details and appearance of the mammogram images.

Regarding the third parameter, the proposed method does not require a training stage of the machine learning algorithms like methods (Dhungel [4], Kai Hu [C4], Ojeda-magana [A7]). Meaning that, the proposed method does not build previously classifiers for detecting the cancer masses in the mammogram images. In contrast, other methods (i.e. Sheng-chih Yang [A6], Kashyap [A3]) do not support the machine learning algorithms or neural networks. However, the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

proposed method distinguishes itself from these methods by running without the need of the initial parameters as shown in the fourth parameter. Thus, the proposed method is able to be more applicable with different types of mammogram images than these compared methods.

The fifth parameter shows the clustering strategy used to detect the cancer mass. These strategies are PSCSM (Progressive Support-Pixel Correlation Statistical Method),BSIF & LBP (Binarized Statistical Image Features & local binary pattern) and PFCM (Possibilistic Fuzzy c-Means clustering algorithm). Here, the Mahalanobis distance has been used as a clustering strategy. It distinguishes itself from these strategies through low computational complexity and emphasizing the high accuracy of detection.

Finally, the running time has been only given in [15]. The size of the mammogram images used in [15] is less than the size used in this study. However, the proposed method has been detected the cancer mass (in approximately 1.2 second) faster than method in [15].

6. CONCLUSION

Within this work, the detection of the cancer mass has been addressed in terms of the breast cancer and mammogram images. A proposed method has been therefore developed for improving the detection results of the cancer mass of the breast cancer. The improvement has been addressed in reducing the running time and low computation complexity. To achieve this improvement, the proposed method has been run in two stages. Firstly, the foreground and background area have been automatically separated in order to avoid the training stage in the detection process. In second stage, the Mahalanobis distance value has been developed to smoothly extract the cancer mass from the foreground area. The proposed method has been developed as a medical application called "cancer cell classification". The latter has been uploaded as a pilot application in the project VI-SEEM for more testing by experts. Moreover, the proposed method has been tested in a dataset of mammogram images. The dataset consist of many of mammogram images taken by CC and MLO views for patient breast. Some of these images are for negative case (diagnosed as breast cancer) and other are for free cases. In addition, these images are for both sides of breast. The proposed method distinguishes itself from other method by automatically running (without initial parameters and training stage). The proposed method reduces the computation complexity by using Mahalanobis distance value. The detection results are improved in term of performance as shown in the detection outputs and running time. The latter is approximately 1.2 second. The limitation of this study is to run the proposed method for detection the cancer mass from given mammogram images. The detection does not include the diagnosis process and features extracting for the cancer mass.

In future, the proposed method will be developed to match both CC and MLO view for mammogram images and to extract features required for diagnosis.

REFRENCES:

[1] A. Chattaraj and A. Das, "Mammographie image segmentation using kernel based FCM clustering approach," in 2016 International Conference on Computer, Electrical Communication Engineering (ICCECE), 2016, pp. 1–6.

[2] M. F. Atoum and Y. M. Al-Khatib, "Association between serum 25-hydroxy vitamin D concentration and taqi vitamin D receptor gene polymorphism among jordanian females with breast cancer," Chin. Med. J. (Engl.), vol. 130, no. 9, pp. 1074–1078, 2017.

[3] A. Shrivastava, A. Chaudhary, D. Kulshreshtha, V. P. Singh, and R. Srivastava, "Automated digital mammogram segmentation using Dispersed Region Growing and Sliding Window Algorithm," in *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, 2017, pp. 366–370.

[4] V. Bhateja, M. Misra, and S. Urooj, "Human visual system based unsharp masking for enhancement of mammographic images," *J. Comput. Sci.*, vol. 21, no. Supplement C, pp. 387–393, Jul. 2017.

[5] GONZALEZ, *DIGITAL IMAGE PROCESSING USING MATLAB 2E*. Tata McGraw-Hill Education, 2009.

[6] J. Anitha and J. D. Peter, "A spatial fuzzy based level set method for mammogram mass segmentation," in *2015 2nd International Conference on Electronics and Communication Systems (ICECS)*, 2015, pp. 1–6.

[7] N. Dhungel, G. Carneiro, and A. P. Bradley, "Fully automated classification of mammograms using deep residual neural

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

networks," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 310–314.

[8] R. Platania, S. Shams, S. Yang, J. Zhang, K. Lee, and S.-J. Park, "Automated Breast Cancer Diagnosis Using Deep Learning and Region of Interest Detection (BC-DROID)," in *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology,and Health Informatics*, New York, NY, USA, 2017, pp. 536–543.

[9] A. Akselrod-Ballin, L. Karlinsky, S. Alpert, S. Hashoul, R. Ben-Ari, and E. Barkan, "A CNN based method for automatic mass detection and classification in mammograms," *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, vol. 0, no. 0, pp. 1–8, Jul. 2017.

[10] G. R. Jothilakshmi and A. Raaza, "Effective detection of mass abnormalities and its classification using multi-SVM classifier with digital mammogram images," in *2017 International Conference on Computer, Communication and Signal Processing (ICCCSP)*, 2017, pp. 1–6.

[11] J. Lian, B. Shi, M. Li, Z. Nan, and Y. Ma, "An automatic segmentation method of a parameter-adaptive PCNN for medical images," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 9, pp. 1511–1519, Sep. 2017.

[12] Y. Jiao, J. Wu, and L. Jiao, "An image segmentation method based on network clustering model," *Phys. Stat. Mech. Its Appl.*, vol. 490, no. Supplement C, pp. 1532–1542, Jan. 2018.

[13] J. Sharma, J. K. Rai, and R. P. Tewari, "A combined watershed segmentation approach using k-means clustering for mammograms," in *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, 2015, pp. 109–113.

[14] H. Min, S. S. Chandra, N. Dhungel, S. Crozier, and A. P. Bradley, "Multi-scale mass segmentation for mammograms via cascaded random forests," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, 2017, pp. 113–117.

[15] K. L. Kashyap, M. K. Bajpai, and P. Khanna, "Globally supported radial basis function based collocation method for evolution of level set in mass segmentation using mammograms," *Comput. Biol. Med.*, vol. 87, pp. 22–37, Aug. 2017.

[16] B. Gupta and M. Tiwari, "A tool supported approach for brightness preserving contrast enhancement and mass segmentation of mammogram images using histogram modified grey relational analysis," *Multidimens. Syst. Signal Process.*, vol. 28, no. 4, pp. 1549–1567, Oct. 2017.

[17] S. Yan, Y. Wang, F. Aghaei, Y. Qiu, and B. Zheng, "Applying a new bilateral mammographic density segmentation method to improve accuracy of breast cancer risk prediction," *Int. J. Comput. Assist. Radiol. Surg.*, Jul. 2017.

[18] S.-C. Yang, "A robust approach for subject segmentation of medical Images: Illustration with mammograms and breast magnetic resonance images," *Comput. Electr. Eng.*, vol. 62, no. Supplement C, pp. 151–165, Aug. 2017.

[19] B. Ojeda-Magaña, J. Quintanilla-Domínguez, R. Ruelas, L. G. Barba, and D. Andina, "Improvement of the Image Sub-Segmentation for Identification and Differentiation of Atypical Regions," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 32, no. 01, p. 1860011, Jul. 2017.

[20] M. Alfraheed, A. Alamouri, and S. Jeschke, "A MDV - Based approach for appearance enhancement of historical images," in *Automation, Communication and Cybernetics in Science and Engineering 2011/2012*, Springer, Berlin, Heidelberg, 2013, pp. 545–557.

[21] The Free Dictionary by Farlex, "craniocaudal," *The Free Dictionary*. [Online]. Available: https://medical-dictionary.thefreedictionary.com/craniocaudal. [Accessed: 19-Apr-2018].

[22] The Free Dictionary by Farlex, "mediolateral oblique view," TheFreeDictionary*.com*. [Online]. Available: https://medical-dictionary.thefreedictionary.com/mediolateral+oblique+view. [Accessed: 19-Apr-2018].

[23] "CCC (classification of Cancer Cells)," *GitLab*. [Online]. Available: https://code.vi-seem.eu/salman.matalgah/CCC/tree/master. [Accessed: 22-Feb-2018].