International Conference on Information and Communication Technologies (ICICT 2014)

A Novel Approach for Removal of pectoral muscles in Digital Mammogram

Sreedevi S[a],*, Elizabeth Sherly[b]

aDept of Computer Science, University of Kerala,Trivandrum,695034
bIndian Institute of Information and Teccknology Management – Kerala, Technopark, Trivandrum,695581

Abstract

Removal of noise and pectoral muscles are the two important pre-processing steps in CAD system for the diagnosis of breast cancer. This work combines Robust Outlyingness Ratio (ROR) mechanism with extended NL-Means (ROR-NLM) filter based on Discrete Cosine Transform (DCT) for the detection and removal of noise. This method removes Gaussian and impulse noise very effectively without any loss of desired data. For segmenting and removing pectoral muscles, this paper uses global thresholding to identify pectoral muscles, edge detection processes to identify the edge of the full breast and connected component labelling to identify and remove the connected pixels outside the breast region. The result shows that our approach removes Gaussian and impulse noise effectively without any loss of desired data and overall gives 90.06% accuracy.

Keywords:.Digital Mammogram,; Breast Region Extraction,; Pectoral Muscle Extraction; Discrete cosine transform; Non local Mean , Medio-Lateral Oblique (MLO) view .

1. Introduction

Breast Cancer is one of the main reasons of death, in middle-aged women since it is very difficult to identify. It cannot prevent completely since its cause remains unknown. But the detection of breast cancer at is earlier stage

* *Corresponding author. Tel.: +91 9447071271;.*
E-mail address:sathyansree123@gmail.com

helps to decrease the death rate and better treatment options available to patients. The cure rate of breast cancer can be increased if it is identified at its earlier stage [1]. This earlier stage detection can be achieved with mammography. Mammograms cannot prevent breast cancer, but they can save the life of a cancer patient by finding breast cancer as early as possible. Mammography helps for the detection of intangible tumors and there by increases survival rate.

Radiologists examine these mammogram images for the detection of abnormalities and to identify whether it is benign or malignant. But sometimes the radiologist fails to differentiate between false positive and false negative. So at times the correct detection of abnormality is impossible by a human radiologist. But the development of Computer-Aided-Detection (CAD) systems helps to assist radiologists in the detection of breast cancer at its earlier stage with great accuracy and stability of diagnosis [2]. The major aim of CAD is to improve the sensitivity by assisting radiologists to detect the suspicious lesion. Though a variety of CAD systems for breast cancer detection is available, further developments are required in this field to produce more effective result.

To detect suspicious lesion from digitized mammogram automatically, some preprocessing steps has to be done for enhancing the quality of the image and for making it ready for further processing . The quality can be improved by removing the unwanted regions in the background of the mammogram image. Removal of noise and pectoral muscles are the two important pre-processing steps in CAD system for the identification of breast cancer. The objective of removing noise from a mammogram image is to get the best estimation from the noisy data for further processing. Besides noise, pectoral muscle is another important artifact present in mammogram image which appear similar to mammographic parenchyma that disturbs the correct detection of breast cancer. Henceextracting pectoral muscle from the mammogram image plays a significant role in the detection of lesion successfully. So preprocessing is an indispensable part in the microcalcification detection. Pectoral muscles contain bright pixels which appear as a rectangular region in one side of the mediolateral oblique view (MLO) mammograms as shown in Fig.1
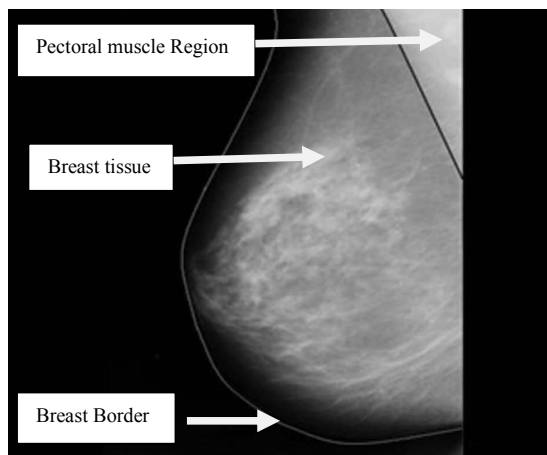

Fig. 1. Anatomy of woman breast.

This study aims to propose an automated method to eliminate two artifacts in a mammogram image. First objective is to introduce a better method for removing noise in the image. Irrespective of the level of noise in each pixel, all the existing methods apply the noise removal algorithm uniformly across the image.This implementation is carried out without checking whether the current pixel is corrupted or not.This would make changes in the intensities of uncorrupted pixels by altering the image details and as a result of this some sort of degradation occurs in the quality of the image. In order to overcome this degradation problem, a new mechanism is introduced to identify corrupted pixels in the image and denoising it , leaving the uncorrupted pixels unaltered. So this paper introduces a new technique formeasuring the level of noise in each pixel andremoving it from the mammogram

without interrupting the uncorrupted portions. The proposed noise removal method is based on the discrete cosine transform (DCT) which is an efficient denoising method and helps to remove all types of noise available in mammogram image. This process has two stages, i.e., detection and then the removal of noise from the image. For measuring the level of noise in each pixel, a statistical measure called ROR is used and for filtering, an extended NL-means are used by introducing a reference image. Then, the proposed method of universal denoising framework uses a combination of ROR with the extended NL-means (ROR-NLM) based on DCT.

Besides noise removal, extracting Pectoral Muscle is also an important operation in Mediolateral Oblique View (MLO). Therefore, the second objective is to identify and remove pectoral muscle in Medio-Lateral Oblique (MLO) view mammograms. In MLO, the pectoral muscles appear slightly brighter than the rest of the breast tissues in the mammogram [3].This work helps to obtain an efficient automatic method to remove noise and to identify the pectoral muscle, which is appears as a rectangular region in one side of the mediolateral oblique view mammograms.

2. Related Works

Today, there are a number of image processing techniques developed by various researchers to identify pectoral muscle in mammogram image. Ferrari et al.[4]introduced an algorithmusing Hough transform for the segmentation of pectoral muscle. The major problem faced by thismethod is the approximation of pectoral muscle by a line. The method does not give desired result if the contour of the pectoral muscles is a curve. To overcome this drawback they introduced Gabor waveletsfor the development of another method[5]which gives better result than the previous one. Kwok et al.[6] proposed Straight line approximation method with edge detection and surface smoothing for pectoral muscle segmentation Kinoshitaet al.[7]introduced the radon domain method which choose the longest straight line for the detection of pectoral muscle. The major drawback of this method is the implementation of a straight line based method on a private database for processing. So this cannot be applicable for common data set. Farag H. Alhsnony et al.[8] propose another technique for the segmentation of digital mammogram image and identifying pectoral muscles using a bit depth and edge processes techniques. Suckling et al.[9] proposed a multiple-linked self-organizing neural network for the segmentation of pectoral muscles. But this method needs a set of good training mammograms to generate satisfactory results. Bajger et al.[10] developed a graph-based approach called minimum spanning tree (MST) method, but this method failed to identify the smaller volume pectoral muscles. Fei Ma et al.[11,12]suggested two graph theory based methods in conjunction with active contours to segment the pectoral muscle in screening mammograms. The two graph theory methods used in their work are Adaptive pyramid (AP) and minimum spanning trees (MST). Camilus etal.[13]proposed a method based on the assumption that mammograms can have curved pectoral muscle edges. In this approach they used graph cut based merging method and a Bezier curve algorithm for pectoral muscle identification. In another approach they used watershed transform method to identify pectoral muscles[14].R. S. Chandra Boss et al propose a method to extract breast region and removal of pectoral muscles using histogram based 8-neighborhood connected component labeling method [15].

3. Proposed Method

The proposed method is represented using the block diagram in Fig.2. To identify the ROI in mammogram, pre-processing must be done to remove noise and pectoral muscles. In the block diagram, the first module represents normalization and the normalized image given as input to the second module which performs noise removal using DCT based NLM filter. Third module represents pectoral muscles identification. Global thresholding is used to identify the pectoral muscles from the mammogram image. After identifying the pectoral muscles, it is removed by the combination of processes of gray level thresholding, canny edge detection and connected component labelling.
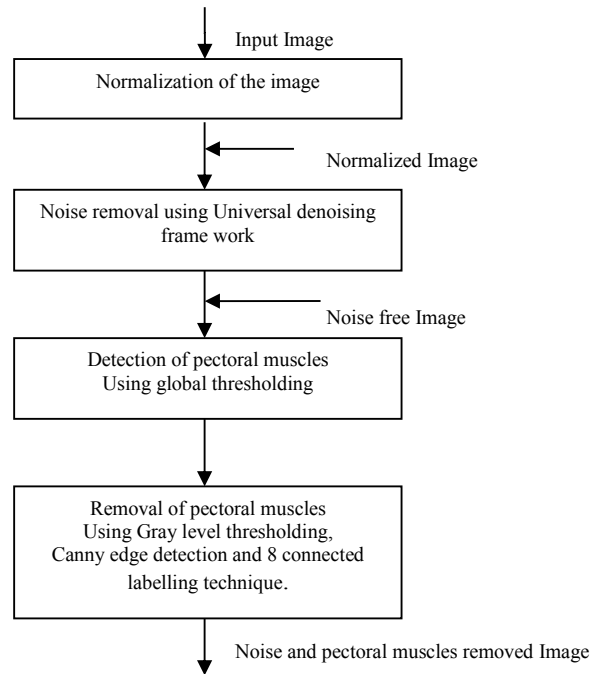
```
                               Input Image
        ┌─────────────────────────────────┐
        │   Normalization of the image     │
        └─────────────────────────────────┘
                          ◄──────  Normalized Image
        ┌─────────────────────────────────┐
        │  Noise removal using Universal   │
        │       denoising frame work       │
        └─────────────────────────────────┘
                          ◄──────  Noise free Image
        ┌─────────────────────────────────┐
        │   Detection of pectoral muscles  │
        │     Using global thresholding    │
        └─────────────────────────────────┘

        ┌─────────────────────────────────┐
        │   Removal of pectoral muscles    │
        │   Using Gray level thresholding, │
        │ Canny edge detection and 8 connected │
        │      labelling technique.        │
        └─────────────────────────────────┘
                   Noise and pectoral muscles removed Image
```

Fig 2. Block diagram of the proposed method

3.1. Normalization

Normalization is the process of bringing the image, or other type of signal, into a range that is more familiar or normal to the senses16. That means it is the process of changing the range of pixel intensity values in an image having poor contrast. To do normalization find oldmax ,oldmin and old range. Old range is calculated using

$$Old\ range = oldmax - oldmin \tag{1}$$

For each pixel find

$$Scale = \frac{(oldvalue - old\min)}{oldrange} \tag{2}$$

$$newvalue = (newrange \times scale) + new\min \tag{3}$$

The gray level normalization formula is given by:

$$g(i,j) = g\min + \frac{(g\max - g\min) \times (I(i,j) - \operatorname{Im}in)}{\operatorname{Im}ax - \operatorname{Im}in} \tag{4}$$

Where Imax and Imin represents the maximum and minimum intensity levels of the original image, gmax and gmin epresents the maximum and minimum intensity levels of the normalized image. g(i, j) is the gray levels at the coordinates of the normalized image and I(i, j) is the gray levels at the coordinates of the original image

3.2. Noise Filtering

Filtering is one of the most important pre-processing steps to eliminate noise present in the image. This paper proposes a frame work for noise reduction in the image. To identify noisy pixels in the image, this work introduces a new statistical measure and proposes a nonlocal means (NL-means) noise-filtering framework based on DCT. DCT based method is more efficient to suppress noise more effectively. Here the similarity weights are calculated in the DCT subspace of neighbourhood instead of computing similarity weights using the gray level information directly. This method is implemented with two stages called detection and filtering. The detection stage calculates the DCT based Robust Outlyingness Ratio (ROR) for measuring the level of noise intensity in each pixel, and then according to the ROR value, all the pixels are divided into four clusters[17]. In the first stage the total number of pixels in the image is divided in to four clusters based on the level of noise intensity in each pixel. Different decision rules are applied in each cluster to detect the impulse noise in the second stage. Then, a new denoising framework is proposed by the combination ROR (new detection mechanism) with the extended NL-means (ROR-NLM) based on DCT. Results obtained by this method shows that the proposed noise detector is superior to other existing detection methods, and the ROR-NLM produces better results and outperforms most existing filters for different noise models. Due to the advantages of DCT, such as low data correlation that helps to reduce inter pixel redundancy and high energy compaction, the DCT based filter is very effective and accurate. The proposed DCT based ROR-NLM also achieves high peak signal-to-noise ratio and great image quality by efficiently removing impulse / Gaussian mixed noise. Thus it will infer a high quality image compared to other filtering techniques.

3.3. Pectoral muscle detection and removal

Pectoral muscles are the triangle shaped region lie in one side, either left or right top corner of mammogram, which contains the brightest pixels in the image. Pectoral muscles appear approximately with similar density as the dense tissues of interest in the image. So identification and removal of pectoral muscles play an important role in detecting the tumor cells. This work proposes a modified tracking algorithm technique for removing pectoral muscle region [18] from the breast region. To avoid the situation of applying different methods for left and right MLO mammograms separately, a general method is designed for detecting pectoral muscle in right MLO mammograms. A left MLO view mammogram is processed by flipping to make it look like a right MLO mammogram.

The filtered image is used for pectoral muscle detection by segmenting the image in to two regions as background region and pectoral muscle region called binarization. For binerization global thresholding and gray level thresholding are used to alter gray image to binary. Here the background region is made completely black and the pectoral muscle region is made completely white. To design a general method for pectoral muscle removal, check whether the mammogram is for the left or right breast. If left it is to be transformed in to the right. In most of mammogram images, canny edge detection method combined with connected component labelling are required to obtain better result. This method allows to suppress the labels and other unwanted artifacts more effectively to produce better ROI.

Algorithm of proposed Method
Pectoral_removal(Img)//Img- denoised Image

Input: Denoised image, Output: pectoral muscle removed image

      Oimg ⟵ Img

      check(Oimg) left or right

      if left
              Img ⟵ fliplr(Oimg)
    end

value ⟵ 1 – global_thresh(Img)
bwimg ⟵ im2bw(Img,value)
bbwimg ⟵ img2bw(Img, graythresh,std)
ebwimg ⟵ edge(bbwimg)
oimg ⟵ bbwimg – bwimg
x ⟵ bwlabel(oimg, 8)
x ⟵ find(x ==1) //breast region extraction
x ⟵ find(x ~=0) // Pectoral muscle removal
.

4. Experimental result

For testing, the proposed method uses mammogram images that are taken from Mammographic Image Analysis Society (MIAS) which is a public database of mammograms.100 images are randomly selected from the normal and abnormal list of images[19] and applied the proposed method on these images. The database contains 322 images of 161 patients. All the mammograms are 8-bit gray level images, digitized at 200 micron pixel edge, and of 1024 × 1024 pixel size.
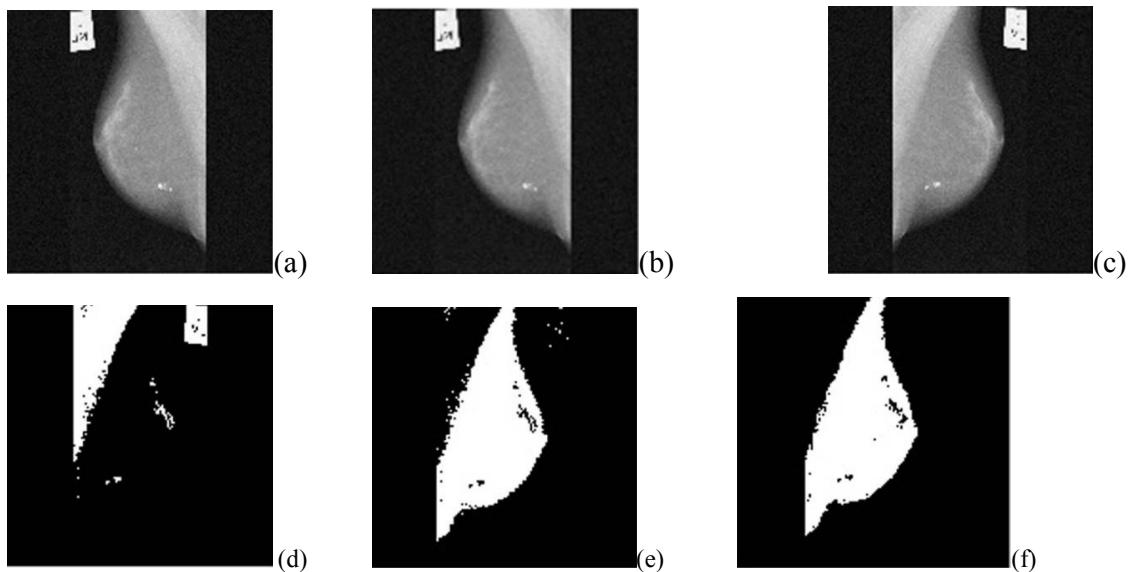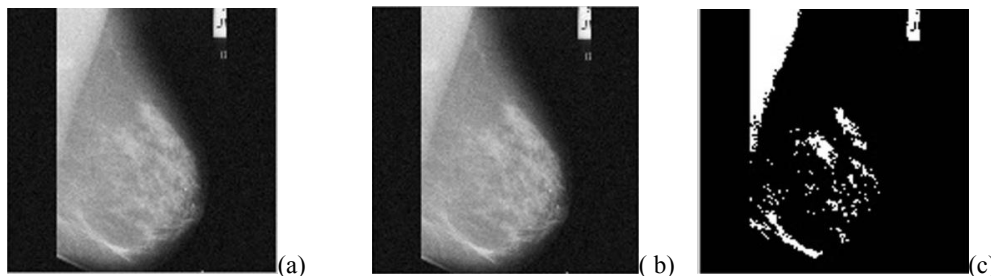

Fig. 3. Result of proposed method.   (a) Original mammogram image (mdb075) (b) noise reduced image (mdb075)(c) Transformed image (mdb075) d) Pectoral muscle identified using global threshold   e) Pectoral muscle removed using global threshold, gray level threshold and canny edge detection technique   f) removed other unwanted information using connected component labelling
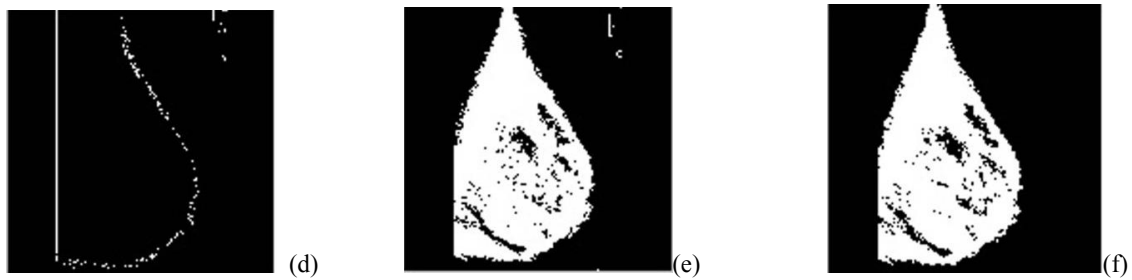
Fig. 4. (a)Original mammogram image (mdb008) (b) noise reduced image (mdb008)   (c)   pectoral muscle identified   using global threshold
 (d) Edge detected using canny operator   (e) pectoral muscle removed using Global threshold, Gray level threshold and edge detection technique
(f) Removed other unwanted information using connected component labelling
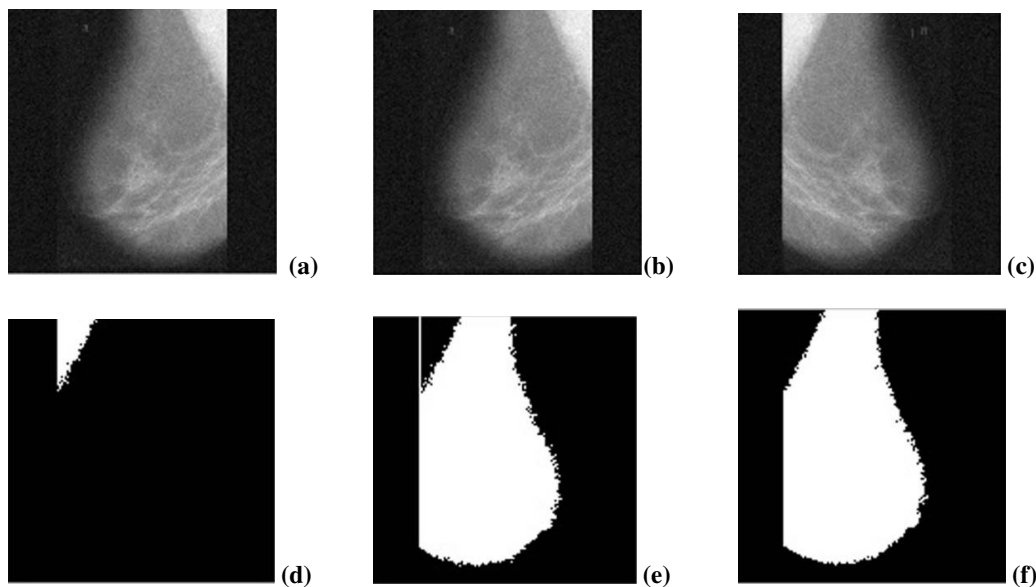

Fig. 5.Result of proposed method. (a) Original mammogram image (mdb027) (b) noise reduced image (mdb027) (c) Transformedimage
(mdb027)   d) Pectoral muscle identified using global threshold   e) Pectoral muscle removed using global threshold, gray level threshold
andcannyedge detection technique   f) Unwanted information removed using connected component labelling

Table 1. Comparative analysis of pectoral muscle removal

Author reference	Methods	No of Images used	Acceptable	Unacceptable
Raba et al[20] , 2005	Adaptive histogram approach	322	278(86%)	44(14%)
R.S. Chandra boss et al[15] 2013	Histogram based 8 neighbourhood connected component labelling	322	288(89.5%)	34(10.5%)
Proposed	Combination of global thresholding, canny edge detection method and connected component labelling	161	145(90.06%)	16(9.94%)

5. Conclusion

In this paper, we applied certain image processing techniques for preprocessing and segmentation of pectoral muscles for further processing of mammogram images. For preprocessing, DCT based non local mean filter is used to remove both Gaussian and impulse noises and then applied canny edge detection technique and connected component labelling after thresholding operations. We experimented with 161 images and the result was compared with similar existing approaches. Our result outperforms and confirmed 90.06% accuracy.

References

1. Ramnath Takiar, Deenu Nadayil, A Nandakumar. Projections of Number of Cancer Cases in India (2010-2020) by Cancer Groups, *Asian Pacific Journal of Cancer Prevention*, 2010; **11**.
2. Wirth MA. Nonrigid approach to medical image registration matching images of the breast, *Ph.D. Thesis, RMIT University*, Melbourne, Australia; 2000.
3. Homer MJ. Mammographic Interpretation: *A Practical Approach*, *McGraw-Hill, Boston*; 1997.
4. Ferrari RJ, Rangayyan RM, Desautels JE, Borges RA, Fr`ere AF. Segmentation of mammograms: identification of the skin-air boundary, pectoral muscle, and fibro-glandular disc, In; *5th International Workshop on Digital Mammography*, 2000; p. 573–**9**.
5. Ferrari RJ, Rangayyan RM, Desautels JE, Borges RA, Fr`ere AF. Automatic identification of the pectoral muscle in mammo-grams,*IEEE Trans on Medical Imaging*, 2004; **23**: p. 232–245.
6. Kwok SM, Chandrasekhar R, Attikiouzel Y. Automatic pec-toral muscle segmentation on mammograms by straight line estimation and cliff detection, in *IEEE Trans on Medical Imaging*, 2004; p.1129–40.
7. Kinoshita SK, Azevedo-Marques P, Pereira RR, Rodrigues JA, Rangayyan RM. Radon-domain detection of the nipple and the pectoral Muscle in mammograms, Journal *of Digital Imaging*, 2008; **21(1)**: p. 37–49.
8. Farag H. Alhsnony, Maher G. M. Abdolrasol, and Samei G. M. Abadelrsool. Auto-Identification of Pectoral Muscle Region in Digital Mammogram Images, *International Journal of e-Education, e-Business, e-Management and e-Learning*, February 2014; **4**: P. 345-67.
9. Suckling , Dance DR, Moskovic E, Lewis DJ, Blacker SG. Segmentation of mammograms using multiple linked self-organizing neural networks, *Med Phys.* 1995;**22(2)**:145–52.
10. Bajger M, Ma F, Bottema MJ. Minimum spanning trees and active contours for identification of the pectoral muscle in screening mammograms. In*: Proceedings in Digital Image Computing, Techniques and Applications. Piscataway*, NJ: IEEE Computer Society Press: 2005.p. 47–53.
11. Ma F, Bajger M, Bottema MJ. Extracting the pectoral muscle in screening mammograms using a graph pyramid*, In: Proceedings APRS Workshop on Digital Image Computing*. Brisbane, Australia: University of Queensland; 2005. p. 39–42.
12. Ma F, Bajger M, Slavotinek JP, Bottema MJ. Two graph theory based methods for identifying the pectoral muscle in mammograms. *Pattern Recogn*. 2007;**40(9)**: p. 2592–602.
13. Camilus KS, Govindan VK, Sathidevi PS. Computer-aided identification of the pectoral muscle in digitized mammograms. *J Digital Imaging*, 2009; **23(5)**: p. 562–80.
14. Camilus KS, Govindan VK, Sathidevi PS. Pectoral muscle identification in mammograms*, Journal of Applied Clinical Medical Physics*,2011:**12(3)**:p.123-34.
15. Subash Chandra Boss R, Thangavel K , Arul Pon Daniel D. Automatic Mammogram image Breast Region Extraction and Removal of Pectoral Muscle, *International Journal of Scientific & Engineering Research*, May-2013:**4(5)**: p.229-35.
16. Thangavel K, Karnan M, Siva Kumar R, Kaja Mohideen A. Automatic detection of microcalcification in mammograms-a review, *Int. J. Graph. Vision Image Process, 2005;***5(5)**:p.31-61.
17.Bo Xiong , Zhouping Yin. A Universal Denoising Framework with a New Impulse Detector and Nonlocal Mean*s, IEEE Transactions on Image Processing,* 2012; **21(4)**:p.45-53.
18. A.Srikarthik, R.Sivakumar, Preprocessing and Enhancement for Mammogram Images Using Unified Approach, *International Journal of Computer & Organization Trends* ,2013;**3(3)**:p.56-61.
19. Suckling J et al.The mammographic image analysis society digital mammogram database, in *Proc. 2nd International Workshop on Digital Mammograph*y, York, England, Elsevier Science, Amsterdam, The Netherlands:1994. p.375-78.
20. Raba David, Oliver Arnau, Mart Joan, Peracaula Marta, Espunya Joan. Breast Segmentation with Pectoral Muscle Suppression on Digital Mammograms*, Springer-Verlag*, Berlin Heidelberg, 2005; p.471–8.