Filtering Method for Pre-processing Mammogram Images for Breast Cancer Detection

Neha N. Ganvir, D. M. Yadav

Abstract: Breast cancer is a stand-out surrounded by the most widely perceived diseases and has a high rate of mortality around the world, significantly risking the health of the females. Among existing all modalities of medical scans, mammography is the most preferred modality for preliminary examination of breast cancer. To assist radiologist, a computer-aided diagnosis (CAD) is enhancing and important medical systems for mammographic lesion analysis. CAD is necessary to provide doctors, to improve detection quality of breast cancer. In mammogram images, micro-calculifications is one of the imperative sign for breast cancer detection. Mammographic medical scan may present unwanted noise and CAD systems are very sensitive to noise. So, pre-processing of medical images for any medical image analysis application like brain tumor detection, breast cancer detection, and interstitial lung disease classification is considered as an important step. The segmentation or classification accuracy is mainly depends upon the significant improved pre-processing process. Thus, in this work, different types of filtering techniques used for noise reduction in medical image processing are analyzed. The qualitative and quantitative results are examined on mini-MIAS mammogram image database. The effectiveness of filtering techniques is compared based on the different quantitative parameters and visual qualities of examined output.

Keywords: Filtering methods, mammogram images, MSE, PSNR, SSIM.

I. INTRODUCTION

The most common inflammation among female across worldwide is breast cancer. The breast cancer is leading factor of deaths for female suffering from cancer diseases. According to the medical survey of India, the breast cancer patients would reach at 1,797,900 by 2020 [6]. This growth of breast cancer patients is because of insufficiency in awareness about health check-up, breast screening, and insufficient medical experts [3]. These factors are very important and necessary to be considered at early stage of breast cancer to confirm prolonged survival. There are different techniques are available to capture breast medical scans namely, computed tomography, mammography, ultrasonography, and magnetic resonance (MRI). Manual annotation of these medical images is a critical and time consuming task. So, for proper diagnosis of the disease robust medical image analysis techniques are required. From mentioned modalities, mammography technique is more suitable and powerful modality for radiologist for preliminary examination of breast cancer [5].

As the screening of mammograms images is challenging and tedious task, the interpretation or clarification about breast cancer is varies from one radiologist to other radiologist using mammogram medical scan. Thus, there is dire need of the expert radiologists to provide more

Figure 1. Basic structure of breast anatomy

Comprehensive diagnosis of breast cancer. To do this task, computer aided detection and a diagnosis (CAD) system is combined with expert radiologists. But, the CAD systems are very sensitive to noise present in the medical images. So, the noise reduction from medical images is very important and key pre-processing step for any medical image analysis applications. In general, different researchers proposed different technique for noise removal using filtering techniques. Thus, in this work, the detailed analysis of existing state-of-the-art methods used for filtering are discussed. The qualitative and quantitative results are examined and compared for detailed analysis.

II. RELATED WORK

A basic tool used to enhance the pictographic representation (texture, shapes, and colors) present in medical/raw data is image processing for human perception and provide more useful and important features for autonomous machine intelligence. Thus, image processing plays very important role in various medical related computer vision applications like medical image analysis. The advancement in image capturing devices and recent technologies like MRI, large amount of data is generated in medical hospital. Because of this growth, there is need to have expert medical services and radiologist in urban as well as in remote places.
One of the leading health issues among women across the world is breast cancer. Early detection of breast cancer is very important for women's health care. Cancer can be diagnosed as the intractable anomalous growth and invasion of cells into different body parts, which forms a mass or lump called as tumor. The tumor can be cancerous or non-cancerous generally known as malignant or benign respectively. The cancer is generally named after the body part where it originates such as lung, breast, prostate, ovarian and thyroid. The epidemiological data of breast cancer depicts incidences and fatality rates by considering various risk factors. The occurrence of breast cancer increases with age, marital status, family history of cancer and stage, menopausal status, pathological nodal status, and educational status. Symptoms of breast cancer are categorized as breast lump, non-lump breast (pain in the breast, anomalies in breast skin or shape and nipples) and non-breast (tiredness, inadequacy of breath, axillary symptoms, neck tumor, and back pain) [15]. Breast comprises of Nipple, Areola, Inflammatory fold, Montgomery’s Glands (Tubercles), Glandular Tissue (lobules), and Retro mammary Space [16].

Understanding anatomical structure of breast is important before proceeding how breast cancer can be detected in its early stage. The breast anatomical structure as shown in Figure 1.

Even before appearance of symptoms, the breast cancer can be detected in its early stages by regular screening tests such as mammography. Local feature based algorithm is proposed for breast cancer detection. The features of breast cancer are assessed through various proposed modalities as shown in Table 1. Various techniques have been developed for detecting microcalcifications in mammogram images. Abirami et al., [17] developed an automatic system for the classification of mammogram images into malignant or benign images utilizing wavelets to extract features and a machine learning approach for the process of classification. The features are extracted using discrete wavelet transforms. These features are the input to the classifier which is based on artificial neural networks for the detection of microcalcifications at each location of the mammograms. Hu et al. [19] developed a microcalcification diagnosis framework which combined the features of HMT as well as wavelet. The DTCWT-HMT effectively modelled the statistical distribution of wavelet coefficients. The relationship between the coefficients of the wavelet also was represented by DTCWT-HMT. After the feature combination process, the features based on DTCWT are extracted and enhanced using GA. These features are given as an input to the ELM classifier for the process of diagnosing malignant and benign clusters of microcalcification. This developed method was compared to various contemporary methods on DDSM, MIAS and Nijmegen datasets with respect to precision and stability proved that the DTCWT-HMT performed the best. A comparison of the proposed method with different classifiers proved that the ELM performed better. The main drawback of this study is that selecting features based on the GA technique is quite slow and also there is a scope for improving the overall accuracy of the classification. Recently, Meenakshi et al. [8] proposed local binary pattern based technique for False-Positive Reduction in Mammograms images as automatic detection of breast cancer system suffers from false positives. In [8], pre-processing (noise reduction) of the mammograms and contrast enhancement is achieved using local textural pattern like binary pattern, ternary pattern, local ternary co-occurrence pattern. Yang et al. [18] developed an automatic method for detecting the clusters of microcalcifications in the digitized mammograms. The area of the breast is determined using Otsu thresholding technique. The contrast of the images is augmented using mathematical morphology. Then the wavelet features are extracted using biorthogonal wavelet. Another approach is proposed [13] for detection of microcalcification in digital mammograms using wavelet features.

### Table 1: Modalities used to measure corresponding Breast features

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Modality</th>
<th>Breast features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mammography</td>
<td>Masses, calcification and other radiographic appearance such as nipple thickening or nipple discharge</td>
</tr>
<tr>
<td>2</td>
<td>Ultrasonography</td>
<td>Guiding interventional procedures such as needle aspirations and localization of breast lump or calcification before breast biopsy</td>
</tr>
<tr>
<td>3</td>
<td>Magnetic Resonance Imaging (MRI)</td>
<td>Provides the important information depicting breast condition, detecting and staging breast cancer</td>
</tr>
<tr>
<td>4</td>
<td>Positron emission tomography (PET)</td>
<td>To detect metastatic disease.</td>
</tr>
<tr>
<td>5</td>
<td>Electrical Impedance Tomography</td>
<td>Based on electrical conductance and capacitance increases as small current pass through the cancer tissues within breast and resulting impedance map highlights malignant area. This is 3D imaging approach.</td>
</tr>
<tr>
<td>6</td>
<td>Thermography</td>
<td>Metabolism and blood vessel proliferation of cancerous tissue increases surface temperature, which is captured using infrared camera forms a high resolution image of these variation.</td>
</tr>
<tr>
<td>7</td>
<td>Galactography</td>
<td>Determines the nipple discharge.</td>
</tr>
<tr>
<td>8</td>
<td>Scintimammography</td>
<td>To determine located lesion in mammogram is malignant. Uses radioactive substance to inject into arm vein and record the accumulation of radiation in the breast.</td>
</tr>
<tr>
<td>9</td>
<td>Computer Tomography (CT)</td>
<td>Adjuvant for monitoring cancer spread.</td>
</tr>
</tbody>
</table>

Published By: Blue Eyes Intelligence Engineering & Sciences Publication

Retrieval Number: A1623109119/2019@BEIESP
DOI: 10.35940/ijeat.A1623.109119

**Filtering Method for Pre-processing Mammogram Images for Breast Cancer Detection**

Abirami et al., [17] developed an automatic system for the classification of mammograms images into malignant or benign clusters of microcalcification. Another approach is proposed [13] for detection of microcalcification in digital mammograms using wavelet features.
The small deposition of the calcium in the breast leads towards breast calcification. For breast cancer detection, most of the methods uses wavelet based features as these feature are able to extract multi-scale features from the image. The dual tree M-band wavelet based features are used for data classifications. Recently, Begum et al. [20] proposed tumor detection and classification approach using wavelet statistical textural features and recurrent neural network. Initially, preprocessing is applied to remove the noise present in the MRI images. After that, textural features are extracted using wavelet transform. Further, these feature are used for classification of tumor with the help of recurrent neural network. Finally, segmentation task is performed using region growing segmentation algorithm.

From literature, it can be observed that different researchers use different algorithms for classifying and detecting the microcalcifications but the accuracy of these techniques is mainly depends on the method which is used for filtering the input image to reduce the effect of noise. Motivation by these facts the main aim of this study is to give the detailed analysis about the different filtering methods used in medical image processing. Out of above discussed modalities, Mammography is the only technique that proves an effective screening technique for the breast [16]. These modalities have certain limitations; effectiveness of ultrasound depends upon radiologist expertise and poor ability to visualize deep lesions; MRI is expensive that requires longer scanning time and contrast agent to be injected for imaging PET does not detect tumors smaller than 5-10mm; EIT has poor signal to noise ratio and has improper modelling of varying impedance between electrode and breast skin; attenuation of infrared radiation in breast tissue is more in Thermography, therefore it provides only surface information of breast; if ducts are incorrectly injected, wrong diagnosis may happen in Galactography; Scintimammography requires injection of substance into human body, which takes longer examination time; CT is less sensitive in detection of microcalcification [16]. Mammography is proven to be most reliable technique for early breast cancer detection and therefore it is ‘gold standard’ for evaluation of breast from imaging modality [8, 17]. The advantages of using mammogram are: At an early stage, detection of breast cancer is helpful in reducing mortality rate with improved medical treatment,
- Can be applied regularly with acceptable side effects,
- Low cost,
- Allow reproducing results

The accurate and speedy interpretation of mammographic screening is affected mainly by two factors i.e. radiologist’s expertise and examination of large volume of cases in the screening program. The deployment of computerized mammographic image analysis and interpretation could improve both accuracy and speed of processing [17]. However, the main hurdle in computerized mammographic image analysis is poor contrast. Preprocessing of mammogram image is performed by removing pectoral muscles and contrast enhancement. The difficulties associated with mammographic images are poor contrast, quantum noise, random fluctuation and variation in X-ray photon absorption in breast tissues that causes difficulty in detecting palpable tumor and leads to more false positive detection. Thus, this research focuses on different filtering methods used for pre-processing the mammogram images. Detail analysis with mathematical expression is given below.

### III. EXISTING PRE-PROCESSING APPROACHES

In this Section, we have discussed about the existing pre-processing approaches like median, wiener, anisotropic, anisotropic filtering with wavelet, etc. The detailed analysis about each filtering method is given below:

#### A. Median Filtering

In general, median filtering is used to remove a salt and pepper noise [10]. It comes into a category of non-linear filtering. This kind of filtering separates the clammers while keeping the sharpness. It ranks the neighboring pixels of a sample pixel according to their brightness level and middle brightness level is considered as a filtered output. The median filtering technique is applied repetitively again and again if required edges present in the images are negligibly humidilate. In case of Mammogram images, some of the researchers [4, 11] make use of a 2D median filtering approach to remove the unwanted straight lines present in the dominant part of mammogram images.

<table>
<thead>
<tr>
<th>Table 2: Existing image enhancement methods</th>
<th>Detailed Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histogram equalization (HE)</td>
<td>Enhances the image by stretching and flattening the histogram of the original image across the complete grayscale range (0–255).</td>
</tr>
<tr>
<td>Contrast Limited Adaptive Histogram Equalization</td>
<td>The noise problem with AHE is minimized by limiting contrast enhancement in homogeneous region i.e. clip the high peak pixel and redistribute them throughout the histogram of an image.</td>
</tr>
<tr>
<td>Histogram Modified Local Contrast Enhancement</td>
<td>It has two stage of operation: 1) histogram modification; 2) local contrast enhancement method. Unlike HE, it provides the adjustment factor to modify level of contrast enhancement.</td>
</tr>
<tr>
<td>Two-dimensional histogram equalization</td>
<td>Uses contextual information, which is obtained from neighborhood pixels to improve visual quality of an image.</td>
</tr>
<tr>
<td>Local gray level S-curve transformation</td>
<td>The overlapping or non-overlapping blocks of fixed dimension is extracted and gray-level S-curve transformation approach is applied on that blocks</td>
</tr>
</tbody>
</table>
Major disadvantage of median filtering is that it largely shuffles the pixel intensities among neighborhoods which might be undesirable in some cases. Also, it may disturb the fine details in the images. Median filtering fails to distinguish between the fine details and the noise.

B. Weiner Filtering

Weiner filtering reduces the overall mean square error (MSE) while processing for noise removal and inverse filtering. The objective of the Wiener filtering is to estimate the unknown signal with the help of associated signal. At the same time, Weiner filtering removes the additive noise and de-blurs the image simultaneously. Weiner filtering is a particular approach for reducing the boisterous parts, in order to improve visual quality of the image [7]. To discover the best recovery of a noisy signal, Wiener filtering is a general method. It provides a straight estimation of noise signal which provides the straight anticipation of the first picture. Based on the comments from the expert Radiologists, Wiener filtering does not affects the structural details of the mammogram image and provides visibility improvement [1]. Researchers make use of Wiener filter to create gauge of a fancied or to target an arbitrary process by LTI sifting of a watched loud process. This clamor filter relies on Fourier cycle. It’s primarily preferred because of the short computational time that it takes to discover the desired solution.

C. Anisotropic Filtering

Initially, Perona and Malik have proposed partial differential equation (PDE) based anisotropic diffusion model for image de-noising. Isotropic diffusion in continuous time domain is given as follows;

$$\frac{\partial M(x,y,t)}{\partial t} = div(\nabla M)$$ (1)

Where, \(M\) is an input mammogram and \(div\) is a mathematical diversion operation. In other way, image smoothing using isotropic diffusion is equivalent to the Gaussian filtering.

$$M(x,y,t+1) = M(x,y,t) + div(\nabla M)$$ (2)

As we know, Gaussian filter is symmetric and rotation invariant which turns into the image blurring. Thus, to avoid this, the anisotropic diffusion can be modified as given below,

$$\frac{\partial M(x,y,t)}{\partial t} = div\left(G(\nabla M)\nabla M\right)$$ (3)

Thus, modified anisotropic diffusion processes only homogeneous regions (non-edgy region) and is does not affects the non-homogeneous region. Image smoothing in terms of modified anisotropic diffusion is defined as,

$$M(x,y,t+1) = M(x,y,t) + div\left(G(\nabla M)\nabla M\right)$$ (4)

The modified anisotropic diffusion process can be discretized as given below:

$$M(x,y,N+1) = M(x,y,N) + \frac{\lambda}{\eta_{(x,y)}} \sum_{(i,j) \in \eta_{(x,y)}} G\left(\left|\nabla M^{(x,y)}_{(i,j)}\right|\right)\left|\nabla M^{(x,y)}_{(i,j)}\right|$$ (5)

where, input mammogram is used as an initial condition \(M(x,y,0)\), \((x,y,0)\) denotes the pixel location, \(N\) denotes the number of iterations, \(\lambda\) is a stability constant <= 0.25, \(\eta_{(x,y)}\) represents the number of neighbourhoods. \(\nabla M^{(x,y)}_{(i,j)}\) represents the pixel difference located at \((x,y)\) and \((i,j)\) i.e. gradient. Four directional gradients can be calculated as given below,

$$\nabla M^{(x,y)}_{((x,y) \rightarrow (x-1,y))} = M(x,y,t) - M(x-1,y,t)$$

$$\nabla M^{(x,y)}_{((x,y) \rightarrow (x+1,y))} = M(x,y,t) - M(x+1,y,t)$$

$$\nabla M^{(x,y)}_{((x,y) \rightarrow (x,y+1))} = M(x,y,t) - M(x,y+1,t)$$

$$\nabla M^{(x,y)}_{((x,y) \rightarrow (x,y-1))} = M(x,y,t) - M(x,y-1,t)$$

Some of the other existing pre-processing/image enhancement techniques are given in Table 2.

D. Anisotropic Filtering With Wavelet

The anisotropic diffusion technique is well known technique used for noise reduction where the edges present in the image is very important. But, one major drawback of the efficient to anisotropic diffusion technique is not efficient when high level noise is present in the image. One major challenge is in medical image processing is that to preserve the information structures present in medical images/scans such as edges and object boundaries to achieve improved performance. Also, the wavelet based transformation gained more importance as able detects the high and low frequency components with multi-scale analysis. Thus, wavelet transform is used in different image processing applications like Smoothing and image de-noising, fingerprint verification, biology for cell membrane recognition, speech recognition, computer graphics and multifractal analysis, etc. To address limitation of anisotropic diffusion and advantage of wavelet transform, wavelet transform domain based anisotropic diffusion technique is proposed for filtering of the medical images. Here, Gaussian smoothing or any other smoothing technique is not used.

$$W^{mn}_{m/2}K(a,b,t+1) = W^{mn}_{m/2}K(a,b,t) + \frac{\lambda}{\sqrt{\sigma}} \sum_{(p,q) \in \Omega(a,b)} g\left([W^{mn}_{m/2}K(a,b,t)],\sigma\right)\left[\left[W^{mn}_{m/2}K(a,b,t)\right]\right]$$ (6)

Where, \(W^{mn}_{m/2}K(a,b,t)\) is wavelet coefficient at time instant or iteration \(t\) at position \((a, b)\), the horizontal and vertical direction is \(m=1,2\), the gradient in DWT domain is \(\Omega\). The threshold used fortuning of gradient magnitude is \(\sigma\), calculated as

$$\sigma = \sqrt{5\sigma}$$ (7)

Where, \(\sigma = \frac{1}{4.2856\psi_k(4\sqrt{k})}\)

Here, the Gaussian smoothing component is removed from the filtering equation. Due to this, this technique is more efficient and more optimal. The gradient is calculated \(W^{mn}_{m/2}K(a,b,t)\) in four different directions as,
The procedure is used for other parameter calculation is same as [2] for anisotropic diffusion algorithm. The anisotropic diffusion is iteratively performed on each scale of wavelet transform components $W_{2i}^{m,n}K(a,b,t+1)$, $W_{2i}^{m,n}K(a,b,t)$ for noise reduction. Here, wavelets based technique is used because wavelet is able to decompose the complex information into elementary forms at different positions and scales. The experimental results are discussed in the next section.

IV. RESULTS AND DISCUSSION

There is need of different appraisal metric qualities to be determined for examine the effectiveness of the proposed filtering method. The extensive experimental analysis is carried out to decide which filtering method is able to give significant improved results for breast cancer detection. For experimental purpose, mini-MIAS database [12] of mammograms database is used. In the medical image analysis application for breast, tumor, lung cancer detection, and input image quality determination is a bottom line and more challenging task for various applications like image enhancement, image restoration, adjusting image quality, etc. Thus, the filtering accuracy is examined in terms of standard deviation (SD), peak signal to noise ratio (PSNR), signal to noise ratio (SNR), self-similarity index measure (SSIM), IMAE, VAR, Root Mean Square Error (RMSE), AMBE. Mathematical formulation of some parameters are given below:

$$SD = \sqrt{\frac{1}{N} \sum (x - \mu)^2},$$  
(9)

where, $x$ is value in data point and $\mu$ is mean

$$SSIM = \frac{2\mu_x\mu_y + c_1(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$  
(10)

$$PSNR = 10\log_{10} \left( \frac{\text{MAX}^2}{\text{MSE}} \right),$$  
(11)

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - K(i,j)]^2,$$  
(12)

$$RMSE = \sqrt{\frac{1}{n} \sum (P - O)^2},$$  
(13)

The quantitative results of the filtered images is examined using different parameter and given in Figure 2. From Figure 2, the anisotropic diffusion with wavelet based filtering is outperformed the other filtering methods in most of the parameters. Here, the advantage of wavelet transform help to extract multi-scale feature related to edges present in the images/medical scan. Also, The qualitative results comparison of different filtering methods on breast Image from mini-MIAS database of mammograms are given in the Figure 3. From qualitative results, it is observed that the filtering method based on anisotropic diffusion filter with wavelet is able to give significant improved edge related information from breast medical images. The parameter values SD, PSNR, SNE, etc. of the proposed filtering method is far better than other filtering methods. Improvements of the results (qualitatively and quantitatively) are evident that the anisotropic diffusion with wavelet filtering is best among the other existing filtering methods.
Filtering Method for Pre-processing Mammogram Images for Breast Cancer Detection

Figure 3: The different filtering methods quantitative results comparison with different evaluation measures on sample images of Benign, Malignant and Normal category from mini-MIAS database of mammograms.

Figure 4: The qualitative results comparison of different filtering methods on breast Image from mini-MIAS database of mammograms (a) Benign Image, (b) Median, (c) Weiner, (d) DWT2 (Haar), (e) Anisotropic Diffusion Filter and (f) Anisotropic Diffusion
filter with wavelet based filtering.

Figure 5: The qualitative results comparison of different filtering methods on breast Image from mini-MIAS database of mammograms (a) Malignant Image, (b) Median, (c) Weiner, (d) DWT2 (Haar), (e) Anisotropic Diffusion Filter and (f) Anisotropic Diffusion filter with wavelet based filtering.

Figure 6: The qualitative results comparison of different filtering methods on breast Image from mini-MIAS database of mammograms (a) Normal Image, (b) Median, (c) Weiner, (d) DWT2 (Haar), (e) Anisotropic Diffusion Filter and (f) Anisotropic Diffusion
Filtering Method for Pre-processing Mammogram Images for Breast Cancer Detection

V. CONCLUSION

Pre-processing of medical images for any medical image analysis application like brain tumor detection, breast cancer detection, interstitial lung disease classification is considered as an important step. The segmentation or classification accuracy is mainly depends upon the significant improved pre-processing process. Also, the mammogram segmentation faces certain challenges such as poor contrast, limited gray tone variations in mammogram images etc. These challenges are overcome by the using pre-processing technique. In this work, different types of filtering techniques used in medical image processing are analyzed. The qualitative and quantitative results are examined on mini-MIAS mammogram image database. The effectiveness of filtering techniques is compared with the help of different state-of-the-art parameters like PSNE, MSE, SSIM, etc. to examine the best filtering methods used to improve the quality of information present in the image. From result analysis, it is evident that proposed anisotropic filtering with wavelet is able to produce significantly improved results (qualitatively and quantitatively) as compared to other filtering methods.

REFERENCES

14. Prasad, N. and D. Houkera, The role of various modalities in breast imaging. Biomedical Papers of the Medical Faculty of Palacky University in Olomouc, 2007. 151(2).

AUTHORS PROFILE

Ganvir Neha N. completed her B. E., M. E., electronics engineering in 2004 and 2010 respectively. Currently she is pursuing her Ph. D under the guidance of Dr. D. M. Yadav from SPPU. INDIA. Her research area is in Image processing, signal processing and biomedical imaging. She is having 15 years of Experience in teaching. She is currently working as assistant professor at SITS Narhe in STES Pune (MS), INDIA.

Dr. Yadav D. M. completed his B. E., M. E. and Ph.D. in electronics engineering in 1994, 2004 and 2009 respectively. His research area is in Image processing, signal processing and biomedical imaging. He is having 24 years of experience out of that 10 years is research. Under his supervision four candidate awarded Ph. D degree and seven are doing their research in Pune University (MS), INDIA. Currently he is Professor and Dean Academic in one of Autonomous College in Pune University.