Automated detection of architectural distortions in mammograms using template matching

O'tega A. Ejofodomi, Edikan Nse Gideon, Gbenga Olalekan Oladipo, Etse Rosemary Oshomah

Department of Electrical and Electronics Engineering, Federal University of Petroleum Resources, Effurun, Nigeria

Email address:

tegae@yahoo.com(O'tega A. E.), Gideon.edikan@gmail.com(E. N. Gideon), lekinsondc@gmail.com(G. O. Oladipo),
pretipanther@yahoo.com(E. R. Oshomah)

Abstract: Breast Cancer is one of the leading causes of death, and its early detection increases the survival rate and treatment options available to patients. Computer-Aided-Detection (CAD) systems have been developed to assist radiologists with the task of locating cancer in mammograms. Unfortunately, these CAD systems have demonstrated less than 50% efficiency in detecting Architectural Distortions (AD), which is a sign of breast cancer. This paper presents a method of detecting AD with better sensitivity results. Forty mammograms containing AD were obtained from the Digital Database for Screening Mammography (DDSM). Each mammogram was preprocessed using breast segmentation techniques to extract the breast region from the mammogram. The mammograms were enhanced using contrast-limited adaptive histogram equalization (CLAHE). Next, the enhanced mammograms were filtered with a bank of 180 Gabor filters to extract the texture orientation from the images. Based on the fact that ADs contain spicules radiating in all direction, AD templates were designed in MATLAB. These templates were cross-correlated with the Gabor filtered mammograms to obtain ROIs that were most likely to contain ADs. The developed algorithm is intended to assist the radiologist by flagging regions likely to contain AD, prompting the radiologist to take a closer look at those regions. The current algorithm developed in MATLAB automatically flags seven suspicious blocks of 25 by 25 pixels per image, and demonstrated a sensitivity of 87.5% with a False Positive per Image (FPI) of 5.2. Future work will focus on the reduction of FPI.

Keywords: Cancer, Architectural Distortion, Gabor Filters, AD Templates, Cross-Correlation

1. Introduction

Breast cancer is one of the leading causes of death. The American Cancer Society estimates that there were 234,580 new cases of breast cancer (232,340 female, 2,240 male) and that 40,030 people (39,620 female, 410 male) died from breast cancer in 2013 [1]. In the United Kingdom, it is estimated that there were 49,961 new cases (49,564 female, 397 male) of breast cancer, and 11, 633 deaths (11,556 female, 77 male) related to breast cancer in 2010 [2]. In some countries, the statistics for breast cancer is unreliable. The cause of breast cancer is unknown, but early detection increases survival rate and treatment options for patients.

The current standard for early detection of breast cancer is screening mammography, which is X-ray imaging of the breast. Typically, radiologists search each image for masses, calcifications and architectural distortions (AD). Architectural distortions are abnormal breast lesions in which the breast parenchyma is distorted with no visible mass. The distortion includes spiculations radiating from a point, and focal retraction or distortion of the edge of the parenchyma [3].

It is estimated that radiologists fail to detect about 10-30% of breast cancers [4-6]. Computer-Aided Detection (CAD) systems have been developed to enable radiologists detect lesions that may be indicative of cancer. These systems serve as support to the radiologist and are not used by themselves as a breast cancer detection tool. CAD systems generally operate by first allowing the radiologist to trace out suspicious areas in the mammogram before the system employs an automated detection algorithm to identify possible cancerous regions. The radiologist then analyzes the regions highlighted by the algorithm to ensure that any cancer present in the image is detected.

Current commercial CAD systems accurately identified 90% cases of masses and microcalcifications [7, 8], but the

sensitivity for AD is low. The R2 Imag e Checker system successfully identified 49% cases containing ADs with 0.7 false positives per image (FPI). The CADx Second Look system [8, 9] successfully detected 33% cases of ADs with 1.27 FPI [7, 9-10]. The efficiency of the current CAD systems in detecting AD is significantly low and emphasizes the need for robust and effective algorithms to improve the sensitivity of the CAD systems. The goal of this work is to develop an algorithm that can assist radiologists in detecting ADs, and also demonstrate better sensitivity than the algorithms employed in the current commercial CAD systems.

Several studies have been conducted on the detection of AD [11]. Guo et al. [12] characterized AD using the Hausdorff fractal dimension, and a support vector machine classifier to distinguish between ROIs exhibiting AD and those with normal mammographic patterns. Tourassi et al. [13] studied the use of fractal dimension to differentiate between normal and architectural distortion patterns in mammographic ROIs. Matsubara et al. [14-16] used mathematical morphology to detect AD around the skin line and a concentration index to detect AD within the mammary gland.

Eltonsy et al. [17] developed a method to detect masses and AD by locating points surrounded by concentric layers of image activity. The sensitivity was found to be 91.3% with 9.1 FPI. Zwiggelear et al. [18] proposed a scheme for the detection of spiculated mass lesions. The abnormal pattern of linear structures and central mass are extracted individually and detection results are combined using principal component analysis (PCA)-based methods. Rangayyan et al. [19] proposed a method based on Gabor filters and phase portrait analysis to detect initial candidates for sites of AD. Rangayyan and Ayres [20] applied Gabor filters to characterize oriented texture patterns and detect AD. The methods were tested with one set of 19 cases of architectural distortion and 41 normal mammograms, and another set of 37 cases of architectural distortion. FROC analysis shows the sensitivity of 0.79 at 8.4 FPI. Sampat and Whitman [21] employed filtering in the Radon transform domain to enhance mammograms, and used radial spiculation filters to detect spiculated lesions. The algorithm was tested on 45 cases exhibiting spiculated masses and on 45 cases with the presence of architectural distortion. A sensitivity of 80% was obtained with 14 FPI in the detection of architectural distortion, and 91% with 12 FPI in the detection of spiculated masses. Zeng et.al. [22] proposed using Gabor filters phase portrait analysis, and feature extraction and classification. Yoshikawa et. al. [23] used adaptive Gabor filters and a concentration index and reported a sensitivity of 82.72%, and 1.39 FPI. Ejofodomi et. al. [24] used a median filter and Gabor filters to extract texture information from 19 images containing AD. AD probability maps were generated using a maximum amplitude map and histogram analysis on the orientation map of the Gabor filter response. AD maps were analyzed to select ROIs as potential AD sites. The map analysis yielded a sensitivity of 79% (15 out of 19 cases of AD were detected)

with 18 FPI.

Advances in the detection of AD in mammograms can better assist the radiologist to improve performance in mammography screening. There is still need for marked improvement in the detection of AD. The algorithm proposed in this paper automatically detects AD by means of Gabor filters and template matching using custom-designed AD templates. The methodology employed is discussed in detail in the subsequent sections, as well as the results obtained using this algorithm.

2. Materials and Methods

2.1. Materials

All images used in this study were obtained from the Digital Database for Screening Mammography (DDSM). A total of 40 images, each containing architectural distortion as identified by a radiologist, were obtained from this publicly available database. The selected images contained both MLO and CC views, had been digitized by different scanners and represented a range of density ratings, subtlety ratings and pathology. The AD algorithm was developed and implemented using MATLAB (The Mathworks™, Natick, MA, USA).

2.2. Methods

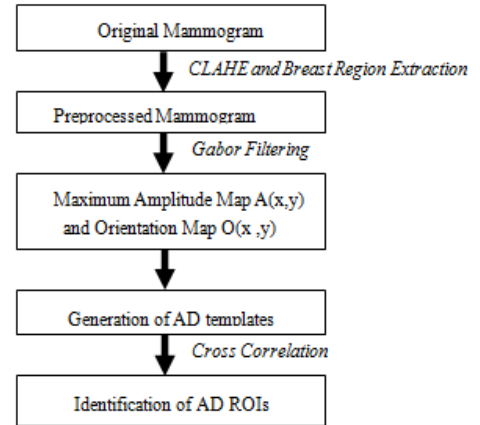A schematic block diagram of the methodology employed is shown in Figure 1. Each step is discussed in detail subsequently.

Figure 1. *Flow Chart for AD Detection.*

2.2.1. Image Preprocessing

Each mammogram was resized and the breast region in the image was extracted and isolated for further processing. This step eliminated the bright label pixels which causes a high false positive detection rate in CAD algorithms [19]. Image Enhancement was performed using contrast-limited adaptive histogram equalization (CLAHE). In CLAHE, the image is divided into small blocks called tiles, and histogram equalization is performed on each block. If any histo-

gram bin is above the specified contrast limit, those pixels are clipped and distributed uniformly to other bins before histogram equalization is applied. After equalization, the neighboring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. A sample of a CLAHE-enhanced mammogram is shown in Figure 2b.
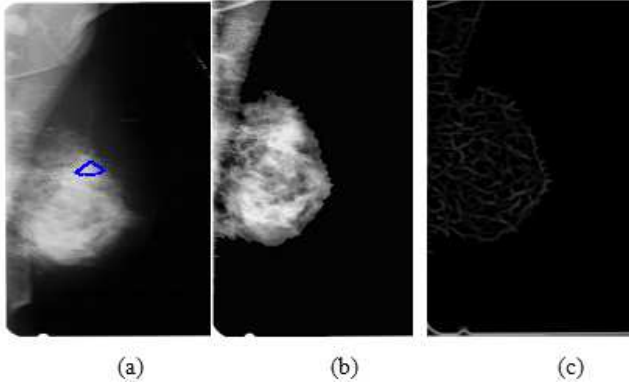

Figure 2. (a) *Original mammogram with radiologist marking (b) breast region segmented and CLAHE enhanced mammogram (c) Maximum output map from Gabor filtration.*

2.2.2. Gabor Filtering

There are many different techniques employed to extract the texture orientation from an image. The method used in this study is the Gabor Filters. The Gabor filter is a sinusoidally modulated Gaussian function. Mathematically, a 2D Gabor function g, is the product of a 2D Gaussian and a complex exponential function. The general expression is given by:

$$g_{\theta,\lambda,\sigma_1,\sigma_2}(x,y) = \exp\left(-\frac{1}{2}\{x\ y\}M\{x\ y\}^T\right)\exp\left\{\frac{j\pi}{\lambda}(x\cos\theta + y\sin\theta)\right\} \quad (1)$$

where $M = \text{diag}(,\sigma_1^{-2}, \sigma_2^{-2})$. The parameter θ represents the orientation, λ is the wavelength, and σ_1 and σ_2 represent scale at orthogonal directions. When the Gaussian part is symmetric, we obtain the isotropic Gabor function:

$$g_{\theta,\lambda,\sigma}(x,y) = \exp(-x^2 + y^2/2\sigma^2)\exp\left\{\frac{j\pi}{\lambda}(x\cos\theta + y\sin\theta)\right\} \quad (2)$$

A bank of Gabor filters was applied to the images. The angle of orientation in the bank of filters was varied from θ ⊡ $[0:\pi]$, the frequency of sinusoidal function was set to $f = 0.088$, and the variance S was set to 0.03. These parameters were chosen experimentally. A total of 180 Gabor filters were applied to the images. Two maps were created from the output of the Gabor filtration: a maximum amplitude map A(x, y) and a dominant orientation map O(x,y). The maximum amplitude for each point (x,y) in the responses to the Gabor filters was extracted to form the amplitude map A(x,y). The angle of orientation of the maximum amplitude was considered to be the dominant angle and was used in creating the orientation map O(x,y).

2.2.3. AD Template Generation

Seven (7) Architectural Distortion templates were designed in MATLAB. The design for each template was based on the prominent feature of architectural distortion being the presence of spicules radiating in all directions [20]. The size of each template was 25 by 25 pixels. Figure 3 shows the seven AD templates used in this study.
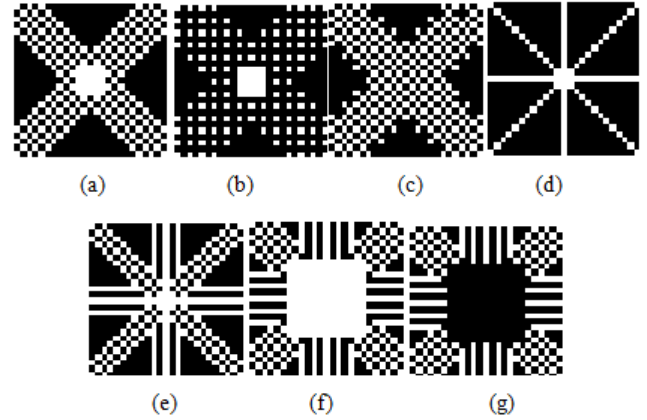

Figure 3. *MATLAB designed AD templates.*

2.2.4. Template Matching using Cross Correlation

A cross correlation algorithm was used to extract regions of interest (ROIs) from the Gabor filtered mammogram. For each template, the cross correlation coefficient between the template and a region of the mammogram was calculated, the size of the region being equivalent to the size of the template i.e. 25 by 25 pixels. In MATLAB, the correlation coefficient was calculated using the following formula:

$$Co(i, j) = \text{corr2}(\text{template}, \text{mammogram}(i{:}i+24, j{:}j+24)) \quad (3)$$

By varying the values of *i* and *j*, different 25 x 25 blocks of pixels in the mammogram are selected for cross correlation computation with the template. The values of *i* and *j* are varied until cross correlation is computed over the entire image. This yields a two-dimensional array containing cross correlation coefficients for each computation, known as the cross correlation map, whose maximum value indicated the location in the mammogram that likely contained AD (see Figure 4). This procedure was repeated for each of the seven templates, resulting in seven locations being flagged as AD in a single mammogram.
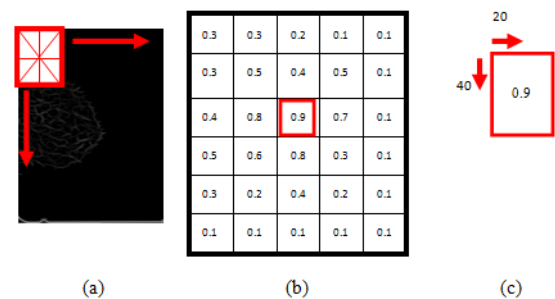

Figure 4. *Cross-Correlation of AD Template and Gabor Filtered Mammogram (a) Cross correlation coefficient between the template and a region of the mammogram was calculated. Process was repeated for subsequent regions resulting in the cross correlation map shown in (b). The location in the map with the highest coefficient identified AD ROI. Row and Column location of AD ROI was obtained in (c).*

2.2.5. Identification of ROIs

The seven locations that are flagged as AD are then highlighted on the original mammogram, to prompt the radiologist to take a closer look at those regions.

Figure 5 shows the output of the algorithm for right MLO and CC mammograms of a single patient. The ground truth, as identified by the radiologist as the true location of the AD, is also shown. A template is considered to have correctly identified AD if its ROI contains pixels that have been identified as being AD by the radiologist's marking (ground truth). If the template's ROI does not contain any of the pixels marked by the radiologist, it has failed to detect AD. In Figure 4a, one template correctly identifies the location of the AD, while the other 6 templates are false positives (FP). In Figure 4b, five templates contain regions of the radiologist's AD marking, while the remaining 2 templates are false positives.
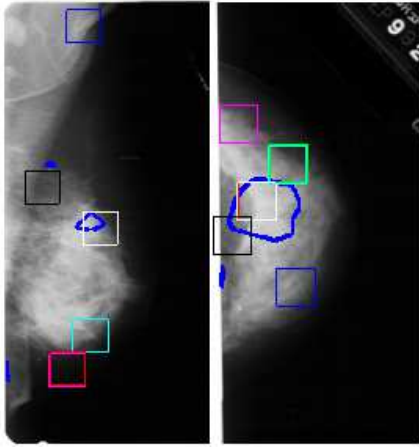

Figure 5. *Output of the AD detection algorithm and radiologists' ground truth overlaid over original mammogram (a) RMLO view (b) RCC view.*

3. Results and Discussion

Table 1 shows the number of true and false positives for the 40 images. Images with 0 TP are those in which the algorithm failed to detect AD. Such images have the maximum number of 7 FP. The overall sensitivity of the algorithm is 87.5% (the algorithm correctly identified AD in 35 of the 40 cases) with an average false positive per image (FPI) of 5.2. FROC analysis of the algorithm is shown in Figure 6.

The developed algorithm is intended to assist the radiologists by flagging regions likely to contain AD, prompting the radiologist to take a closer look at those regions. The current algorithm automatically flags seven suspicious blocks of 25 by 25 pixels per image, and demonstrated a sensitivity of 87.5% with a FPI of 5.2. The FPI can be reduced further by ranking ROIs in terms of certain features, such as fractal dimension. The algorithm can then discard ROIs which do not meet a specific threshold based on the selected features under consideration. Future work will focus on the reduction of FPI.

Table 1. *True Positives (TP) and False Positives (FP) identified in Images.*

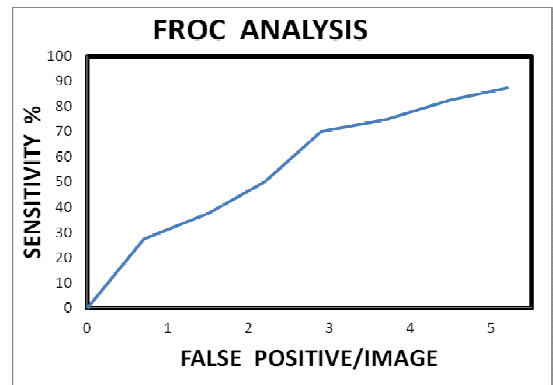| Image No.	Number of True Positives (Tp)	Number of False Postives (Fp)
1	1	6
2	2	5
3	1	6
4	2	5
5	3	4
6	0	7
7	1	6
8	2	5
9	1	6
10	1	6
11	3	4
12	1	6
13	1	6
14	1	6
15	0	7
16	3	4
17	1	6
18	1	6
19	2	5
20	2	5
21	1	6
22	5	2
23	0	7
24	4	3
25	1	6
26	3	4
27	3	4
28	2	5
29	3	4
30	0	7
31	3	4
32	0	7
33	2	5
34	2	5
35	3	4
36	5	2
37	1	6
38	1	6
39	2	5
40	2	5

Figure 6. *FROC Analysis of AD Detection Algorithm.*

There were five images in which the algorithm failed to detect AD. Four of those five images were MLO views. This is likely because of the presence of the pectoral muscle in the MLO view, which shows up as a bright white region. The algorithm often flags regions in the pectoral muscle, particularly at the boundary, and this is a prominent cause of algorithm's false positives and failure to identify AD. Identification and removal of the pectoral muscle in the preprocessing stage could lead to higher sensitivity and lower FPI. Table 2 shows that the algorithm currently has a higher sensitivity than commercial CAD systems.

Table 1. *Comparison of AD Algorithm to Commercial CAD systems.*

	AD Algorithm	Image Checker R_2	CADx Second Look
Sensitivity (%)	87.5	38	21
No. of Images	40	45	45
FPI	5.2	0.7	1.27

The size of the AD templates was selected experimentally. More AD templates could be designed to improve the current sensitivity of the algorithm. However, this will result in an increase in FPI rate, unless feature extraction and classification is incorporated into the algorithm.

4. Conclusion

Current CAD systems have demonstrated less than 50% sensitivity in detecting architectural distortions in mammograms. There is need, therefore, to develop superior algorithms that can assist radiologists in detecting architectural distortions. We have developed an algorithm to detect architectural distortions using MATLAB. The algorithm consists of a preprocessing stage to enhance the image and extract the breast region, a filtering stage to enhance spicules using Gabor filters, and an AD detection stage by comparing the mammogram with custom designed templates using cross correlation. The algorithm was tested on 40 images obtained from the Digital Database for Screening Mammography, and demonstrated a sensitivity of 87.5% (35 out of 40) and a FPI of 5.2. Four of the 5 images were AD went undetected were MLO views. The reason for the algorithm's failure to detect AD is likely because of the presence of the bright white pectoral muscle in the MLO view, as the algorithm often flagged regions in the pectoral muscle, particularly at the boundary. Identification and removal of the pectoral muscle in the preprocessing stage could lead to higher sensitivity and lower FPI. Improvements to the current algorithm can also be made by focusing on the reduction of FPI using feature extraction and classification.

References

[1] American Cancer Society: Cancer Facts and Figures (2013). American Cancer Society, Inc., 2013. www.cancer.org. Accessed February 28, 2013.

[2] Cancer Research, UK. http://www.cancerresearchuk.org. Accessed February 28, 2013.

[3] American College of Radiology ACR BI-RADS - Mammography, Ultrasound & Magnetic Resonance Imaging, 4th edn. Reston, VA (2003).

[4] K. Kerlikowske, P. A. Carney, B. Geller, M. T. Mandelson, S. H. Taplin, K. Malvin, V. Ernster, N. Urban, G. Cutter, R. Rosenberg, R. Ballard-Barbash. "Performance of screening mammography among women with and without a first-degree relative with breast cancer," Annals of Internal Medicine 133: 855-863, 2009.

[5] T. M. Kolb, J. Lichy, J. H. Newhouse. "Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations," Radiology 225: 165-175, 2000.

[6] R. E. Bird, T. W. Wallace, B. C. Yankaskas. "Analysis of cancers missed at screening mammography," Radiology 184: 613-617, 1992.

[7] The CADx Second Look system: http:/www.icadmed.com Accessed January 30, 2014.

[8] The R2 Technology's Image Checker: http://www.r2tech.com Accessed January 30, 2014.

[9] H. Bornefalk. "Estimation and comparison of CAD system performance in clinical settings," Academic Radiology. Vol 12, Issue 6 pp. 687-694, 2005.

[10] M. P. Sampat, M. K. Markey, A. C. Bovik. "Computer-aided detection and diagnosis in mammography," IEEE EMBS pp. 4973-4978, 2004.

[11] A. Kamra, S. Singh, V. K. Jain. "Towards the Detection of Architecture Distortion in Mammograms: A Review. International Journal of Computer Applications," Volume 46 No.7 pp. 44-49, 2012.

[12] Q. Guo, J. Shao, V. Ruiz. "Investigation of support vector machine for the detection of architectural distortion in mammographic images," J. Phys. Conf. Ser., 15: 88–94, 2005.

[13] G. D. Tourassi, D. M. Delong, and C. E. Floyd Jr. "A study on the computerized fractal analysis of architectural distortion in screening mammograms," Physics in Medicine and Biology, 51(5):1299–1312, 2006.

[14] T. Matsubara, D. Fukuoka, N. Yagi, T. Hara, H. Fujita, Y. Inenaga, S. Kasai, A. Kano, T. Endo, and T. Iwase. "Detection method for architectural distortion based on analysis of structure of mammary gland on mammograms," In Proceedings of the 19th International Congress and Exhibition on Computer Assisted Radiology and Surgery (CARS 2005), pages 1036–1040, Berlin, Germany, 2005. Elsevier.

[15] T. Matsubara, T. Ichikawa, T. Hara, H. Fujita, S. Kasai, T. Endo, and T. Iwase, "Automated detection methods for architectural distortions around skinline and within mammary gland on mammograms," In H. U. Lemke, M. W. Vannier, K. Inamura, A. G. Farman, K. Doi, and J. H. C. Reiber, editors, International Congress Series: Proceedings of the 17th International Congress and Exhibition on Computer Assisted Radiology and Surgery, pages 950–955, London, UK, June 2003. Elsevier.

[16] T. Matsubara, D. Yamazaki, T. Hara, H. Fujita, S. Kasai, T. Endo, and T. Iwase. "Automated detection of architectural distortion on mammograms," In H.-O. Peitgen, editor, Digital Mammography IWDM 2002: 6th International Workshop on Digital Mammography, pages 350–352, Bremen, Germany, June 2002. Springer-Verlag

[17] N. Eltonsy, G. Tourasi., A. Elmaghraby. "Investigating performance of a morphological based CAD scheme in detecting architectural distortion in screening mammograms," Proceedings of the 20th International Congress and Exhibition on Computer assisted radiology and Surgery, pp 336-338, 2006.

[18] R. Zwiggelaar, T. C. Parr, J. E. Schumm, I. W. Hutt, C. J. Taylor, S. M. Astley and C. R. M. Boggis. "Model-based detection of spiculated lesions in mammograms," Medical Image Analysis, 3(1), 39-62, 1999.

[19] S. Banik, R. M. Rangayyan, and J. E. Desautels. "Detection of architectural distortion in prior mammograms of interval-cancer cases with neural networks," Proc. IEEE EMBS, 6667-6670, 2009.

[20] R. M. Rangayyan, F. J. Ayres. "Detection of architectural distortion in prior screening mammograms using Gabor filters, phase portraits, fractal dimension, and texture analysis," Int JCARS Vol. 2:347–361, 2008.

[21] M. P. Sampat, G. J. Whitman. "Evidence Based Detection of Spiculated Masses and Architectural Distortions," SPIE. Vol. 5747, pp. 26-37, 2005

[22] X. Zheng, Z. Lao. "Multiscale quantification of tissue spiculation and distortion for detection of architectural distortion and spiculated mass in mammography," Medical Imaging: Computer-Aided Diagnosis, Proc. of SPIE The International Society of Optical Engineering, Vol. 7963, 2011.

[23] R. Yoshikawa, A. Teramoto, T. Matsubara, H. Fujita. "Automated detection scheme of architectural distortion in mammograms using adaptive Gabor filter," Medical Imaging: Computer-Aided Diagnosis, SPIE, Vol. 8670, 2013.

[24] O. Ejofodomi, M. Olawuyi, D. U. Onyishi, G. Ofualagba. "Detecting Architectural Distortion in Mammograms using a Gabor Filtered Probability Map Algorithm," IFIP Advances in Information and Communication Technology, Vol. 412, pp.328-335, 2013.