

Breast Thermograms Asymmetry Analysis using Gabor filters

Priyanka Yadlapalli¹, Reddy Madhavi K², Sunitha Gurram², Avanija J², K Meenakshi¹, Padmavathi Kora¹

¹GRIET, Hyderabad, India

²Sree Vidyanikethan Engineering College, AP, India

Abstract. Women are far more likely than males to acquire breast cancer, and current research indicates that this is entirely avoidable. It is also to blame for higher death rates among younger women compared to older women in nearly all developing nations. Medical imaging modalities are continuously in need of development. A variety of medical techniques have been employed to detect breast cancer in women. The most recent studies support mammography for breast cancer screening, although its sensitivity and specificity remain suboptimal, particularly in individuals with thick breast tissue, such as young women. As a result, alternative modalities, such as thermography, are required. Digital Infrared Thermal Imaging (DITI), as it is known, detects and records temperature changes on the skin's surface. Thermography is well-known for its non-invasive, painless, cost-effective, and high recovery rates, as well as its potential to identify breast cancer at an early stage. Gabor filters are used to extract the textural characteristics of the left and right breasts. Using a support vector machine, the thermograms are then classified as normal or malignant based on textural asymmetry between the breasts (SVM). The accuracy achieved by combining Gabor features with an SVM classifier is around 84.5 percent. The early diagnosis of cancer with thermography enhances the patient's chances of survival significantly since it may detect the disease in its early stages.

Keywords. Breast cancer, Digital Infrared Thermal Imaging (DITI), Gabor filters, edge detection, pattern recognition

1 Introduction

Cancer is the unregulated proliferation of dangerous undesirable or uncontrolled cells, some of which may even overgrow and invade other tissues of the body. Breast cancer is the most common type of cancer in women, out of all cancer types. The sooner cancer is discovered, the better the prospects of recovery and therapy. The human body possesses a high degree of thermal symmetry when it is healthy. Any notable departure from this symmetrical pattern in thermograms is deemed abnormal. As a result, thermograms were analysed quantitatively in automated diagnostic systems to determine the asymmetry between left and right areas [1]. Before extracting any quantitative information, areas of interest were segmented from background tissues. The healthy human body has a high degree of thermal detection of cancer in its early stages because vascularity (angiogenesis) in the breast changes as a result of temperature pattern. Localized or distributed heat patterns can be treated as benign breast lesions. [2] Thermography can identify tumours with an average size of 1.28 cm, but mammography detects tumours with an average size of 1.66 cm, implying that cancer can be identified at an earlier stage using thermograms than mammography. The method described in this study extracts textural characteristics from images to identify asymmetry between bilateral breasts caused by tumours. In this study, the

texture characteristics are based on Gabor filters. The extracted features are classified using Support Vector Machine (SVM) to classify the thermograms into healthy and cancerous. For many years, different works on breast cancer detection have been reported. Usha Rani Gogoi et al. [3] focused mostly on statistical feature extraction to distinguish normal thermograms from malignant thermograms. The classification approach is discussed in Bartosz Krawczyk and Gerald Schaefer [4], who examine the topic of class distribution in medical pattern analysis. The method for determining the ROI, segmentation, and asymmetry analysis of breast thermograms is described in [5]. The characteristics derived from thermograms are based on the temperature profile of the breast surface. These characteristics give information on the pathophysiology of breast thermograms. H. Ghayoumi Zadeh et al. and may be utilised as an intelligent communicative system to speed up and improve the accuracy of diagnosis.

2 Methodology

The recommended approach for detecting normal and abnormal instances in thermograms is designed in four key phases. These stages are as follows [10-16]:

- A. Data collection
- B. Comparison of normal and atypical breast asymmetry
- C. Thermogram feature extraction by extracting textural features using gabor filters

D. SVM classifier classification

An outline of the proposed system is shown in Fig.1, all the modules are explained in detail below

A. Data Collection

The thermal imaging data is not immediately accessible, either online or otherwise. The breast thermograms in this study were obtained from the DMR (Database for Mastology Research) Database [11]. This information is most often utilised in ongoing research for breast cancer early detection. These photos were captured with the FLIR SC-620 Thermal Camera. The Hospital of the Federal University of Pernambuco donated the pictures utilised in the tests. In this work, 35 thermograms of asymmetric breast thermograms and 35 thermograms of healthy breast thermograms are used for training and classification. In thermal imaging, there are different types of acquisition techniques. These methods can be static or dynamic, as well as single, sequential, or accompanied capture. When a patient's body temperature matches that of the surrounding environment, the body is subjected to thermal stress[6]

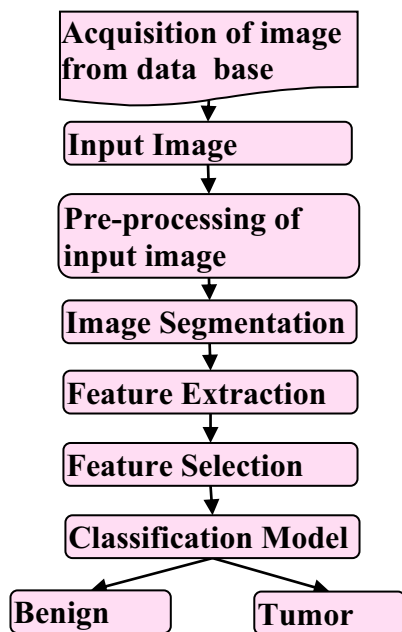


Fig.1. Proposed system

B. Asymmetry study between normal and abnormal breast

Because the thermal symmetry of a healthy human body is so great, any changes from the original symmetry may be deemed abnormal. As a result, thermograms are used to detect these anomalies. Any departures from the original symmetry may be regarded as abnormal. As a result, thermograms are used to detect these anomalies

a) Pre-processing

Before further processing, regions such as the neck, chest, and other portions of the breast thermogram must be eliminated. As a result, the backdrop and extraneous region of the breast thermogram are manually

eliminated. The thermograms are then transformed into grayscale pictures.

b) Extraction of a Region of Interest and Segmentation

The thermograms as they are obtained from a closer view, the thermal image that has been acquired has a maximum active area that is protected by the breast and is regarded as the active region, whilst the framework of the thermal image is regarded as the passive zone of the thermal image. As all unnecessary sections of the breast thermograms are manually eliminated, segmentation plays a critical role in the total study of breast asymmetry. The technique described in this work not only discards the background region, but also extracts the left and right breasts using horizontal edge detection. Otsu's thresholding and Morphological procedures are used to detect horizontal edges. The procedures are as follows:

Unwanted/unnecessary parts are removed from breast thermograms manually.

- The thermograms are converted into binary pictures using Otsu's thresholding.
- After converting thermogram pictures to binary images, morphological procedures are used to patch gaps.
- The majority of the ones in the binary picture may be discovered by utilising the labelling idea.
- The vertices of largest part are used to remove background from the thermograms.
- By extracting the horizontal edge the lowest breast edge can be found
- The image is cropped from the lowest breast edge
- The image is split in to two images from the centroid.

These steps are shown in the Fig. 2.

C. Feature extraction

After image segmentation is done for identifying normal thermograms vs aberrant thermograms, or for classification purposes, features extraction is critical. In this study, textured features are retrieved using the Gabor filter.

To characterize the textural characteristics, a multiresolution feature extraction technique is utilized. These characteristics are defined at multiple sizes and orientations in order to differentiate between different retrieved features. This is a Gaussian modulated sinusoid with multiresolution breakdown in the spatial and spatial frequency domains. Bovik et al [17] presented a model for filter placement based on individual texture power spectrum properties. A dyadic Gabor filter bank is utilised in the work of Jain and Farrokhnia to analyse the spatial-frequency domain [18]. While all prior experiments have shown that the Gabor filter is effective at identifying both texture frequency and orientation.

The real impulse response of a 2-D sinusoidal plane wave with orientation and radial centre frequency

modulated by a Gaussian envelope with standard deviations – and along the x and y axes is given by

$$g(x, y) = \frac{1}{2\pi\rho_x\rho_y} \left\{ -\frac{1}{2} \left[\frac{x^2}{\rho_x^2} + \frac{y^2}{\rho_y^2} \right] \right\} \quad (1)$$

where $x = \cos \theta + \sin \theta$ $y = -x \sin \theta + y \cos \theta$

In the equivalent spatial-frequency domain, the Gabor filter might be represented by two symmetrically spaced Gaussians, as shown below.

$$g(x,y) = \exp\{-2\pi^2[(u-f)^2\sigma_x^2 + v^2\sigma_y^2]\} + \exp\{[(u-f)^2\sigma_x^2 + v^2\sigma_y^2]\}$$

(2)

After setting the frequency cut-off to -6 db and the frequency and orientation bandwidths to constant values that match psycho visual data, the values of x and y are computed using the equation below. Experiments revealed that the frequency bandwidth of simple cells in the visual cortex is about one octave. A circular, Gaussian was chosen for this study by choosing $x = y$ to have equal spatial coverage in all directions and a 45° orientation bandwidth.

$$\sigma_x = \frac{\sqrt{\ln 2}(2^B+1)}{\sqrt{2}\pi(2^n-1)} \quad \sigma_y = \frac{\sqrt{\ln 2}}{\sqrt{2}\pi f \tan \frac{B_n}{2}}$$

(3)

Setting the filter properties carefully would result in good texture capture and decrease the influence of aliasing. This is accomplished by setting the filter position (f_0) and bandwidth (x, y) appropriately, as well as ensuring that the centre frequencies of channel filters are close to typical texture frequencies to avoid the filter response from dropping off too quickly.

The mean was initially removed from each of the 560 420 pictures used in this study to minimise the filter's sensitivity to texture with continuous fluctuation

Then four radial frequencies of 0.25, 0.176, 0.125, 0.088 A total of 20 filters were used, with five orientations of 0°, 36°, 72°, 108°, and 144°.

Finally, the recovered features would reflect each filter's magnitude response, also known as Gabor energy. The following equations are used to compute a bank's orientations (O) and frequencies (F).

$$O = \frac{(j-1)\pi}{v} \quad \text{Where } j=1 \text{ to } 5$$

(4)

$$F = \frac{f_{max}}{\sqrt{2^{l-1}}} \quad \text{Where } i=1 \text{ to } 5$$

(5)

The process of applying the collection of filter banks is similar to the operation of wavelet modification of a picture at chosen spatial frequencies. As is the case with translation and scaling of the mother wavelet basis generation, the Gaussian function is modulated and translated for the creation of the Gabor basis functions. The Gabor function, on the other hand, is regarded an appropriate wavelet since the basis created by the Gabor function is non- orthogonal, resulting in repetitive decompositions. The specification of Gabor filter banks is crucial since it predicts the number of necessary radial frequencies based on picture size.

The classical method for extraction of Gabor filter texture signature is the energy E_k , where $k=1, 2$ as it is shown in the below equation where M and N are the size of the sub band intensity

$$E_k = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [I(x, y)]^k \quad (6)$$

Following the creation of the filter banks, each segmented picture is convolved with a filter bank that has five orientations and four frequency shifts, yielding a total of 20 features (absolute mean or energy). As a result, all 35 aberrant and 35 normal segmented pictures were obtained. Gabor features are extracted, and a set of 40x70 features (40x35 normal and 40x35 aberrant) is fed into a classifier for classification.

3 Classification

After extracting valuable textural features, by thermogram SVM classifier classifies those thermograms into normal and abnormal(benign) . SVM is a classification technique which takes input as labelled data from two classes and outputs a newly classified unlabelled or labelled data into one of the two classes. SVM employs two fundamental processes necessary for classification, namely training and testing. SVM is trained with data that is familiar and has recognised class values. The SVM is then trained to categorise unknown data using its training set.

To apply an SVM classifier using a linear or nonlinear mapping function, the input vectors are mapped to higher dimensional features, and a separate hyperplane is constructed that is at the greatest distance from the closest points. A training set is then created for the application of SVM for classification. For linear separable data, the linear decision function $g(x)$ is as follows.

$$g(y) = w^T \cdot y + c \quad (7)$$

Where w and c stand for the weight and bias parameters, respectively. For each training sample x_i , $g(x_i) > 0$ when $z_i = 1$ and $g(x_i) < 0$ when $z_i = -1$. The hyperplane is then used to divide training samples from two distinct classes.

$$g(y) = w^T \cdot y + c = 0 \quad (8)$$

The SVM classifier uses a hyperplane to divide the two groups thereby , maximizing the separation distance between them .

4.Results and Discussions

In this work, 35 thermograms of healthy patients and 35 pictures of malignant patients with a resolution of 640 x 480 pixels are obtained from the DMR data set. Frontal and side views are available in both static and dynamic acquisitions. This research was conducted using both static and frontal thermograms. Figures 3 and 4 exhibit normal and malignant pictures from the data set Figures 4 and 5 illustrate healthy and aberrant pre-processed breast thermograms with unnecessary portions eliminated, as well as segmented thermograms.

All 35 aberrant and 35 normal segmented pictures have their Gabor characteristics retrieved. There are 20 gabor characteristics in each picture. As a result, a collection of 40x70 features (each from the left and right breasts) is submitted to an SVM classifier for categorization.

Due to ontogenesis, or the development of anatomical features from the earliest portion, the bilaterals (left and right breasts) of a malignant patient have distinct temperature profiles. This asymmetry between the two breasts occurs in abnormal instances, whereas symmetrical breast thermograms are found in patients with no malignant development.

The classifier is critical in the breast cancer detection system because it divides thermograms into normal and abnormal groups. The base SVM classifier is trained and tested with 40 features from each thermogram, 35 normal and 35 abnormal thermograms.

This method achieves an average accuracy of 85 percent and a maximum accuracy of 91.5 percent. The precision, sensitivity, and specificity are calculated as follows:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (9)$$

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

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

The accuracy, sensitivity and specificity of the six iterations are given in the Table no.1.

Table 1. Classification Results using SVM

ITERATION S	SENSITIVIT Y (%)	SPECIFICIT Y (%)	ACCURAC Y (%)
1	92.51	87.84	90.5
2	81.62	83.08	82.84
3	91	78.8	81.39
4	74.96	76.34	76.12
5	100	84.6	87.89
6	100	82.16	87.18

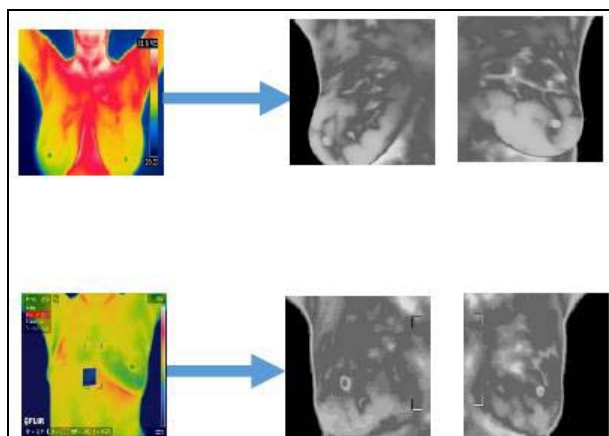


Fig. 2. Segmented Normal Thermograms

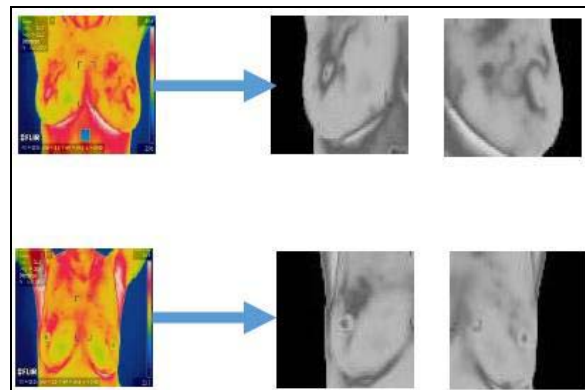


Fig. 3. Segmented Abnormal Thermograms

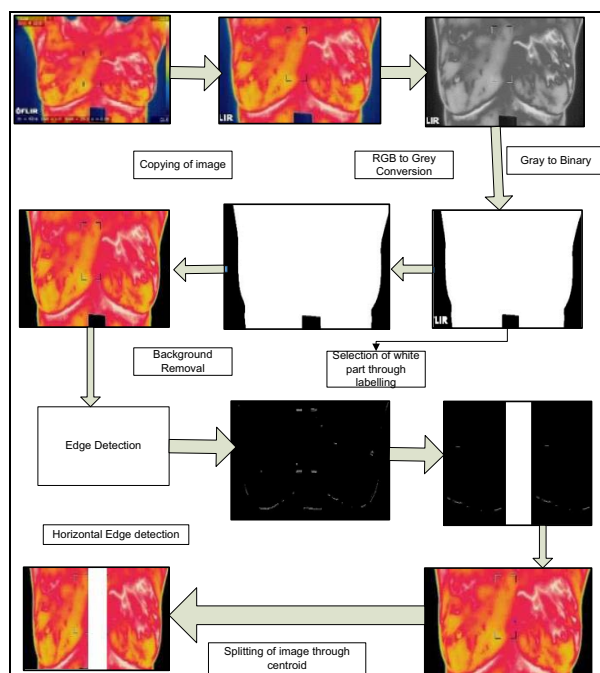


Fig. 4. Extracting ROI and Segmentation

The average level of accuracy achieved is 85.4 percent, with a maximum level of accuracy of 92.15 percent. The high sensitivity numbers represent a small number of missed instances of breast cancer. For the asymmetry analysis the segmentation of ROI is very important. The normal and abnormal instances are then categorized using the SVM classifier, which has an accuracy of 85.4 percent and high sensitivity and specificity values.

5. Conclusion

Breast cancer is a major health problem for women all over the world. Breast cancer patients have a high mortality rate. The mortality rate can be lowered to a greater extent by early detection through appropriate and effective screening. Thermography has emerged as a formidable competitor to other screening modalities like as mammography. Breast cancer is a major health problem for women all over the world. Breast cancer patients have a high mortality rate. The mortality rate

can be lowered to a greater extent by early detection through appropriate and effective screening. Thermography has emerged as a formidable competitor for other screening modalities such as mammography, as the latter identifies tumours while their size is considerably smaller as compared to mammography, owing to increased blood flow in the tumour and ontogenesis. When compared to other procedures, this approach is far superior since it is less expensive, non-invasive, non-ionizing, and portable. Women in rural and remote areas can take this screening tests due to its portability. In this work, breast cancer detection using infrared thermal imaging is explored with good accuracy of 85.4%.

References

1. S. Malvia, S. A Bagadi, U. S. Dubey and S. Saxena, *Asia- Pacific Journal of Clinical Oncology* (2017); 13: 289–295.
2. Krawczyk, Bartosz, and Gerald Schaefer. *Systems Journal*, IEEE 8.3 (2014): 921-928.
3. Wakankar, Asmita, G. R. Suresh, and Akshata Ghugare. *Electronic Systems, Signal Processing and Computing Technologies (ICESC)*, 2014 International Conference on. IEEE, (2014).
4. Zadeh, H. Ghayoumi, et al. *Canadian Journal on Image processing and computer vision* 2.6 (2011):54-58.
5. L.F. Silva, D.C.M. Saade, G.O.S.O. Silva, A.C. Paiva, R.S. Bravo and A. Conci, *Journal of Medical Imaging and Health Informatics*, vol 4(1),pp 92-100(9),(2014).
6. Arora, Nimmi, et al, *The American Journal of Surgery*, 196.4 (2008): 523-526.
7. Neela, Tejashwini, and Swetha Namburu. *IET Networks* 10.3 (2021): 103-109.
8. Sandeep Chaplot et al, , *Biomedical Signal Processing and Control* 1, pp. 86–92, (2006).
9. Hajare, Pravin S., and Vaibhav V. Dixit. *International Journal of advancement in electronics and computer engineering (IJAECE)*, pp. 116-119,(2012).
10. Kuraparathi, Swaraja, Meenakshi Kollati, and Padmavathi Kora. *Traitement du Signal*, 36, 6, (2019).
11. Kora, Padmavathi, and K. Sri Rama Krishna. *Information systems design and intelligent applications*. (Springer, New Delhi, 2016), pp. 309-317.
12. Kora, Padmavathi, K. Meenakshi, K. Swaraja, A. Rajani, and Md Kafiul Islam, *Informatics in Medicine Unlocked*, 17, (2019), pp. 100257.
13. Swaraja, K, J. *Adv. Res. Dyn. Control Syst.(JARDCS)*, 11, (2017).
14. Swaraja, K., Y. Madhaveelatha, and V. S. K. Reddy, *Int J Appl Eng Res*, 11, 1, (2016), pp. 216-225.
15. Kora, P., & Kalva, S. R, *SpringerPlus*, 4, 1, (2015), pp. 1-19.
16. Dhanalaxmi, B., G. Apparao Naidu, and K. Anuradha. *Procedia Computer Science*, 46, (2015), pp. 432-442.
17. A. C. Bovik, M. Clark, and W. S. Geisler, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55–73, (1990).
18. A. K. Jain and F. Farrokhnia, *Pattern Recognition*, vol. 24, no. 12, pp. 1667–1686, (1991).

* Corresponding author: priyankayp@gmail.com