Local and Global thresholding-based breast cancer detection using thermograms

Vartika Mishra*[1], Subhendu Rath[2] and Santanu Kumar Rath[1]

[1] NIT Rourkela, Rourkela, Odisha, India, 769008
[2] VCU, Richmond, VA, USA 23219

vartikamishra151@gmail.com, subhendu.rath@vcuhealth.org, skrath@nitrkl.ac.in

Abstract. In a recent study, to detect breast cancer abnormalities, thermography has been observed to be a qualitative modality. Due to an increase in blood vessel activity, the cancer cells and the tissues become hotter. Thus, by using thermography, the thermograms of the image are captured which depict the surface temperature of the breast, where the hot spot is indicated by the higher temperature. This modality has increased the non-invasiveness in detecting breast abnormalities. The different channels of the RGB (Red, Green, and Blue) image gives information of the different intensity of the color with respect to the image. In this work, the breast thermograms corresponding to the red channel are extracted for analysis. Further, three different thresholding methods viz., Otsu thresholding, Adaptive Mean Thresholding, and Adaptive Gaussian Thresholding methods are applied which depicts the local image characteristics of the image. The image enhancement methods improve the quality of the image. Further, the statistical features are extracted from the obtained thresholded images, and two different classifiers Random Forest and Decision Tree are applied for classifying the normal and abnormal breast.

Keywords: Thermography, color channels, thresholding, feature selection

1. Introduction

Cancer is one of the second most cause of global death accounting 9.6 million of death as recorded by World Health organization [1]. One of the most common cancers among different cancers is breast cancer due to which 6,27,000 deaths have been reported. Hence, the early detection and treatment of the cancer can subsequently reduce the mortality rate. Breast cancer is usually detected in ducts, glands producing milk, tubes carrying milk to the lobules and nipples. Different modalities have been helpful in the detection of the breast cancer viz. Ultrasound, Mammography, Magnetic Resonance Imaging (MRI), thermography and many more.

Mammography is one commonly adopted methodologies in which the images are captured with X-ray exposures compressing the breasts. It has been observed that, difficulty arises when capturing the dense breast tissues thereby, the tumors of the respective area leaving undetected. Infrared Thermography is one of such techniques that has the potential to detect breast cancer even to the extent of eight to ten years earlier than diagnosis by application of Mammography [2]. IR is a non-ionizing, radiation free and non-contact technique. It helps in detecting the tumors based on the asymmetry analysis when both left and right breasts are compared for the tumor detection [3].

In this work, the red channel of the RGB image is taken for analysing the breast thermograms which helps to identify the abnormality of the breast thermograms. The red channel image is extracted from the RGB images. Further, these obtained red channel images are converted to the grey scale images. Adaptive mean thresholding, Adapting Gaussian thresholding and Otsu thresholding is applied for the breast thermograms for differentiating the background from the image and to get more precise edges. Further, the breast image is cropped into left and right part of the breast. The different statistical features are extracted from these breast thermogram images viz. Gray Level Co-Occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Neighbourhood Gray Tone Difference Matrix (NGTDM), Gray Level Size Zone Matrix (GLSZM) and Gray Level Dependence Matrix (GLDM). Relief-F methods is applied for feature selection method and further Decision Tree and Random Forest are applied for classifying among the healthy and unhealthy breast thermograms.

2. Literature Survey

This section presents a review of articles on breast thermogram classification based on different methods of feature extraction methods, feature selection methods and their analysis. Motta et al. proposed a method for tri-dimensional profile based on the temperature profile by applying the level set method for the automatic segmentation of the breast thermograms [4]. They applied Otsu thresholding for background segmentation. Ritam et al., proposed super pixel-based segmentation for automatic segmenting the breast thermograms [5].

Gogoi et al., have obtained a feature set of 24 features by applying the Mann Whitney Wilcoxon test. They applied six different classifiers among which ANN and SVM with RBF kernel gave the highest accuracy or 84.75% [6]. D. Sathish et al., have extracted the texture features in spatial domain. Further, by applying SVM classifier and ANN classifier, the later is giving a better accuracy of 80% [7]. De Santana et al., have extracted the geometry and texture features viz. Zernike and Haralick moments. Among different classifiers applied Extreme Machine Learning (ELM) and Multi-Layer Perceptron (MLP) obtained promising results. An accuracy of 76.01% is attained with Kappa index [8]. Sathish et al., have proposed a novel method based on temperature normalization of the breast thermograms and extracted the wavelet based local energy features. Further RSFS and GA are applied for selecting the best set of features among which RSFS is selecting the best set of features when classified with SVM classifier with Gaussian kernel function giving an accuracy of 91% [9]. R. Karthiga et al., have extracted the spatial domain and curvelet domain based statistical features and applied hypothesis testing. 16 best set of features are selected and classified with different classifiers viz. Logistic Regression, SVM, K- Nearest Neighbor, Naïve Bayes. Among these above-mentioned classifiers, cubic- SVM is giving a highest accuracy of 93%. [10]. A comparative analysis of the state of the art is represented in a tabular form in Table 1.

Table 1. State of the art comparison

Author's Name	Dataset	Methodology Applied
Motta et al. [4]	DMR database	Level Set Method
Ritam et al. [5]	147 thermograms	Super-pixel segmentation
Gogoi et al. [6]	80 thermograms	Mann Whitney test
Sathish et al. [9]	100 thermograms	Temperature based normalization
Karthiga et al. [10]	60 thermograms	Extracted spatial and curvelet domain features

It is observed from the above-mentioned literature survey that better preprocessing of images further helps in extracting better set of features. Thus, this research gap motivates towards getting a better vision of the images. The proposed work is based on the RGB image which corresponds to the information of different color channels for the respective image. The red channel images are extracted which indicates the higher probability of tumor for abnormal breast. Further, three thresholding methods are applied viz. Otsu thresholding, Adaptive Mean thresholding and Adaptive Gaussian thresholding. Texture features are extracted from these images and Relief-F method is applied for feature selection. The best set of features are classified with Decision Tree and Random Forest.

3. Dataset

For this study, a publicly available dataset comprising of a total number of 56 breast thermograms is considered. This dataset is available online at the Database for Mastology Research (DMR) repository of UFF, Brazil [11]. The breast thermograms, are captured with the FLIR SC-620 THERMAL camera having specification of 640 X 480 spatial resolution which has 37 unhealthy and 19 healthy breast thermograms.

4. Breast thermogram analysis

In this study, breast thermograms are obtained by transforming the temperature matrix into images. On these obtained images, the red channel images are extracted from the three channel RGB images. Further, they are converted to gray scale and thresholding-based techniques are applied for the images. Three different methods viz., Otsu thresholding, Adaptive Mean thresholding and Adaptive Gaussian Thresholding is applied for obtaining the object from the background. Further, different texture features are extracted by application of GLCM, GLSZM, GLRLM, NGTDM and GLDM. After extracting the features from thermograms, feature selection method Relief-F is applied for the selection of the most distinctive features. The feature subset obtained are then classified for healthy and unhealthy subject. The respective flow of the work is explained in the form of the flow chart as shown in Fig. 1.

4.1. Pre-processing

The breast thermograms are obtained by transforming the temperature matrix. The images obtained are in RGB channel, where the red channel images are extracted for the analysis of the abnormality of the breast thermograms. Further, different thresholding methods, adaptive thresholding and Otsu thresholding methods are applied for the images obtained.

Thresholding methods: The thresholding methods are divided into local thresholding and global thresholding. The traditional thresholding method was proposed wherein a suitable single threshold value was chosen to distinguish between the background and the object [12]. The pixel values above the threshold value are classified in one class and the pixel values below the threshold value are classified in another class. Selecting an optimal threshold value is a very important key for applying the thresholding technique. An Image I (a, b) of N levels of gray in an image gy (a, b) is defined with two gray levels as:

$$gy\,(a,b) = \begin{cases} 1, & if\ I(a,b) \geq T \\ 0, & if\ I(a,b) < T \end{cases} \quad ...(1)$$

where pixels corresponding to value 1 are objects and the pixels corresponding to 0 are background.

4.1.1. Adaptive thresholding:

The adaptive thresholding method is a local method where the threshold values of smaller regions are calculated. For finding the local threshold, the intensity values of each pixel are statistically examined for local neighborhood. It calculates different threshold for different regions of the image and gives results which are better with varying illumination [13].

a. **Adaptive Mean thresholding**
The value of threshold is calculated by the mean of neighbouring (a, b) pixels.

b. **Adaptive Gaussian thresholding**
The value of threshold is calculated as the gaussian weighted sum of block size neighbourhood of (a, b) pixels.

4.1.2. Otsu Thresholding:

It is a global thresholding technique and is applied on the bimodal images. It chooses the threshold value based on the minimization of the within class variance by minimizing the weighted sum of class variance [14].

$$\sigma_w^2(t) = q_1(t)\,\sigma_1^2(t) + q_2(t)\,\sigma_2^2(t) \quad ...(2)$$

$$\sigma_b^2(t) = q_1(t)[1 - q_1(t)][\mu_1(t) - \mu_2(t)]\ \hat{}\ 2 \ ...(3)$$

$$\sigma^2 = \sigma_w^2(t) + \sigma_b^2(t) \ldots (4)$$

where,
σ^2 : $the\ total\ variance,$
$\sigma_w^2(t)$: $the\ within\ class\ variance$
$\sigma_b^2(t)$: $the\ \mathrm{b}etween\ class\ variance$

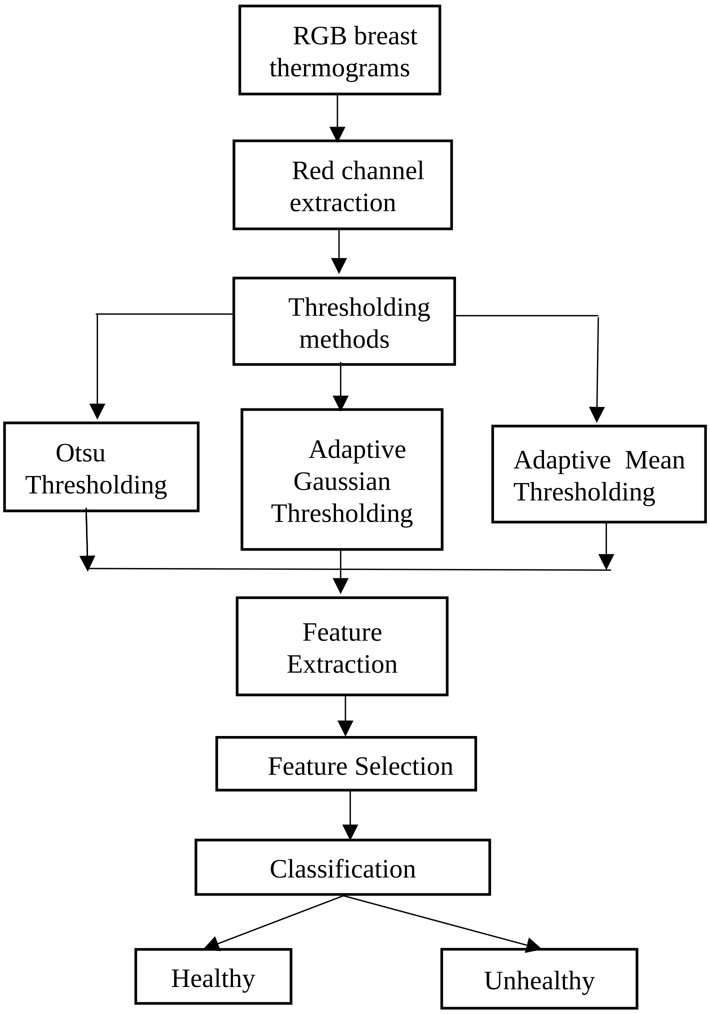

Fig. 1 Structure of the experiment executed

The images are further segmented between left and right breast using the manual segmentation method as shown in Fig. 2.

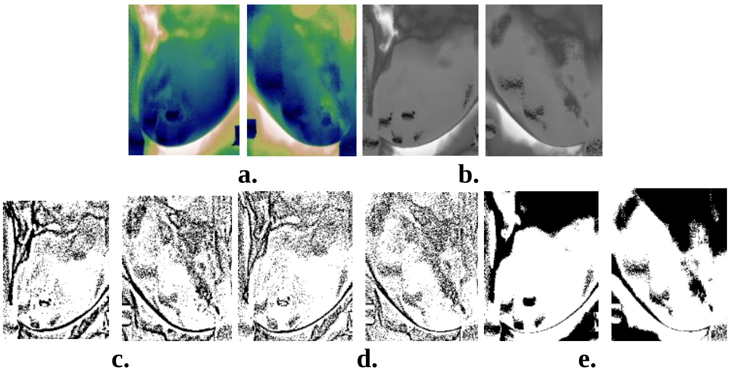

Fig. 2. The segmented left and right breast thermograms: a. the breast thermograms; b. the red channel breast thermograms; c. the Adaptive Mean thresholding; d. the Adaptive Gaussian thresholding; e. the Otsu thresholding.

4.2. Feature Extraction

The different techniques of feature extraction play a very significant role in detecting the breast cancer abnormalities. It is observed that the texture patterns are better reconginzed by the gray level intensity in the thermogram in the particular direction. The different intensity levels describe the mutual relationship among the neighbouring pixels of the image. In this work, from the segmented left and right breast, various features are extracted by applying techniques such as GLRLM, GLCM, GLSZM, GLDM and NGTDM.

4.2.1. Gray Level Co- Occurrence Matrix (GLCM)

GLCM is a second order statistical feature extraction method which helps to analyze the textural properties of the image [15]. The different gray levels in the image mark the spatial orientations adjacent to the pixels from each other. Twenty four different features are extracted from the breast thermograms viz. joint average (JA), autocorrelation (AC), cluster shade (CS), difference entropy(DEN), cluster tendency(CT), correlation (CR), inverse variance(IV), sum of squares (SOS), informational measure of correlation 2 (IMC2), contrast(C), inverse difference moment (IDM), maximum probability (MP) inverse difference normalized (IDN), inverse difference moment normalized(IDMN), joint entropy (JEN), sum entropy (SEN), inverse difference (ID), sum average (SA) difference average (DA), difference variance (DV), joint energy (JEY), maximal correlation coefficient (MCC), cluster prominence (CP) and informational measure of correlation 1 (IMC1),

4.2.2. Gray Level Run Length Matrix (GLRLM)

GLRLM statistical features are computed by the length of number of consecutive pixels which has same gray level value. Here the occurrence of given gray colour is calculated on the basis of direction giving homogeneous pixels running for every gray-levels [16]. Total sixteen GLRLM features viz., long run emphasis (LRE), gray level non-uniformity normalized (GLNUN), short run emphasis (SRE), run length non-uniformity normalized (RLNUN), gray level variance (GLV), long run high gray level emphasis (LRHGLE), run percentage (RP), run entropy (REN), short run low gray level emphasis (SRLGLE), gray level non-uniformity (GLNU), run length non-uniformity (RLNU), high gray level run emphasis (HGLRE), low gray level run emphasis (LGLRE), short run high gray level emphasis (SRHGLE), run variance and long run low gray level emphasis (LRLGLE) are extracted from the breast thermograms.

4.2.3. Gray Level Size Zone Matrix (GLSZM)

GLSZM feature extraction method quantifies the different gray level zones in an image [17]. Her, the connected pixels share the same gray level intensity. The features extracted from the matrix are large area emphasis (LAE), small area emphasis (SAE), size zone non-uniformity (SZNU), gray level non-uniformity normalized (GLNUN), zone variance (ZV), zone percentage (ZP), size zone non uniformity normalized (SZNUN), low gray level zone emphasis (LGLZE), small area high gray level emphasis (SAHGLE), small area low gray level emphasis (SALGLZE), gray level non-uniformity (GLNU), large area high gray level zone emphasis (LAHGLE), large area low gray level emphasis (LALGLE), zone entropy (ZEN), high gray level zone emphasis (HGLZE) and gray level variance (GLV).

4.2.4. Gray Level Dependence Matrix (GLDM)

GLDM statistical texture features quantifies gray level dependencies [18]. Here, the features are calculated as the number of pixels which are connected within are depending on the centre pixel. The different features extracted from the matrix are Large Dependence Emphasis (LDE), Dependence Non-Uniformity (DNU), gray Level Non-Uniformity (GLNU),), Dependence Entropy (DEN), Dependence Variance (DV), Low gray Level Emphasis (LGLE), Dependence Non-Uniformity Normalized (DNUN), Large Dependence High gray Level Emphasis (LDHGLE), High gray Level Emphasis (HGLE), Small Dependence High gray Level Emphasis (SDHGLE), gray Level Variance (GLV), Small Dependence Emphasis (SDE), Small Dependence Low gray Level Emphasis (SDLGLE) and Large Dependence Low gray Level Emphasis (LDLGLE).

4.2.5. Neighbourhood Gray Tone Difference Matrix (NGTDM)

NGTDM texture features are calculated with quantifying difference based on the different gray level values and their average gray value which are within the distance [7]. The different features extracted are Coarseness (Co), Busyness (B), Contrast (C), Strength (SH) and Complexity (CL).

4.3. Feature Selection

The feature selection methods help in removal of the redundant and irrelevant features from the extracted set of features. This further assists in better result analysis with the selected subset of features. In this work, Relief-F is applied for selecting the best set of features.

The main idea behind implementing the Relief-F method is it helps in estimating the most qualitative features. It calculates on the basis of the distinguishes between different instances which are near to each other. Relief-F method finds the two nearest neighbors, each from the different class. This results in the adjustment of the feature weighting vector, where more weight is given to the features that are discriminating between the instances from the neighbors, belonging to different class [19].

4.4. Classification

A classification model is used to draw conclusion from the input instances given for training. It predicts the class for the new instance given for training. The classifier maps the given instance to that specific category. In this study, two different classifiers Decision Tree [20] and Random Forest [21] are applied for classifying the healthy and unhealthy breast thermograms.

5. Results Analysis

In this study, the red channel-based images are extracted from the RGB image. Further the three thresholding methods viz. Otsu thresholding, Adaptive Mean thresholding and Adaptive Gaussian thresholding methods are applied for distinguishing the background from the object. Further, the thresholded images are applied for extracting the texture features viz. GLCM, GLRLM, GLDM, NGTDM and GLSZM. Further, two classifiers are applied for classifying the breast thermogram analysis among healthy and unhealthy breast.

Table 2. The average PSNR metrics calculated for different thresholding methods for healthy and unhealthy breast

Thresholding Methods	Normal	Abnormal
Otsu Thresholding	55.63	55.06
Adaptive Mean Thresholding	56.61	55.05
Adaptive Gaussian Thresholding	58.61	58.05

The PSNR (peak signal to noise ratio) value for the reconstructed image is calculated for the different three thresholding methods and it is observed that higher the value of the PSNR, the better is the quality of the image. The PSNR value differentiates the similarities between the original image and the reconstructed image. It is also defined as the root mean square of each pixel in the image. The different values for the image obtained are shown in Table 2, where the average values of the image are calculated. It is observed that the Adaptive Gaussian thresholding is giving the highest value thus depicting to be a better technique for reconstructing the image.

It is calculated as:

$$\text{PSNR} = 20 \; * \; log_{10}(max\ value / \ RMSE) \quad ...(5)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{rm}\sum_{j=1}^{cm}(I_o\ (a,b) - I_r\ (\ a,b))}{r_m * r_c}} \quad ...(6)$$

where,

$r_m \; * \; r_c$ are the maximum number of rows and columns, $I_o\ is\ the\ original\ image$, I_r is the reconstructed image respectively?

The Otsu thresholding method is applied when the grey level intensities are clearly distinguished in the image. In this work, Otsu thresholding method is applied for the red channel extracted images which depicts the higher temperature of the surface of the breast thermograms. The features are extracted from these images and further applied for classification. Random Forest and Decision Tree are applied for classifying among the healthy and unhealthy breast thermograms. Among the both classifiers, Random Forest is giving a better accuracy of 92.30% as compared to the decision Tree giving an accuracy of 85.90%. The parameters viz., precision, sensitivity and specificity are also calculated and are shown in Table 4. The positive predictive value for random Forest is high as the precision is giving a higher value of 93.39% as compared to the Decision Tree.

Table 3. Comparison of state of the art with the proposed work

Author's Name	Classifier	Accuracy
Gogoi et al. [6]	ANN	84.29%
Sathish et al. [7]	ANN	80%
De Santanna et al.[8]	MLP	76.01%
Sathish et al. [9]	SVM	91%
Karthiga et al. [10]	SVM	93%
Proposed Work	Random Forest	94.67%

Table 4. The performance metrics in percentage obtained for the Otsu Thresholding images

Classifiers	Accuracy	Sensitivity	Specificity	Precision
Random Forest	92.30	93.39	95.07	86.96
Decision Tree	85.90	90.22	83.46	89.02

The Adaptive thresholding method is applied on the red channel of the breast thermograms which helps in analyzing the abnormality in the abnormal breast as the surface temperature of the unhealthy breast is higher due to the presence of the tumor in the cells. Block size of 11 x 11 is taken for the image for calculating the threshold value of every pixel block wise. Further, the features are extracted from these obtained images and two classifiers viz. Decision Tree and Random Forest are applied for classifying between the healthy and unhealthy breast thermograms. Among the two classifiers, Random Forest is giving a better accuracy of 94.67% by applying the Adaptive Gaussian thresholding method. Another method Adaptive Mean thresholding is giving an accuracy of 91. 12% which is comparatively less. The other performance parameter values are also calculated and are shown in Table 5 and Table 6 for each thresholding method based on different classifiers.

Table 5. The performance metrics in percentage obtained for the Adaptive Gaussian Thresholding images

Classifiers	Accuracy	Sensitivity	Specificity	Precision
Random Forest	94.67	93.02	98.52	88.89
Decision Tree	88.46	90.45	91.71	82.40

Table 6. The performance metrics in percentage obtained for the Adaptive Mean Thresholding images

Classifiers	Accuracy	Sensitivity	Specificity	Precision
Random Forest	91.12	92.96	91.56	88.00
Decision Tree	89.35	93.52	90.18	87.72

However, among the three thresholding methods, images applied by Adaptive Gaussian thresholding are giving a better accuracy of 94.67 % when Random Forest is applied for classifying the unhealthy and unhealthy breast as compared to the decision tree. The positive predictive value is higher for random forest classifier than decision tree as shown in Table 5. A box plot is calculated shown in Fig. 3. describing the accuracies obtained for different thresholding methods viz. Otsu thresholding, Adaptive Mean thresholding and Adaptive Gaussian thresholding for two different classifiers viz. Random Forest and Decision Tree. A comparative observation of the proposed work with the state of the art is discussed in Table 3.

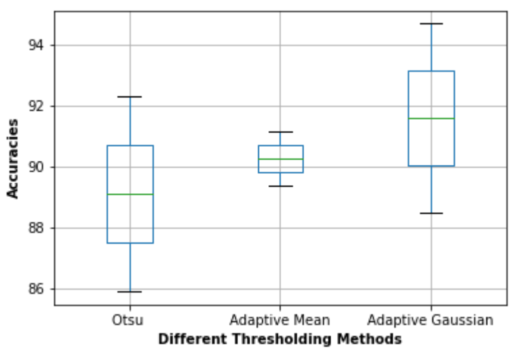

Fig. 3 Box plot of the accuracies for the different thresholding methods

6. Conclusion

In this study, the red channel-based images are extracted from the RGB images. Further, three thresholding methods viz. Otsu thresholding, Adaptive Mean thresholding, and Adaptive Gaussian thresholding methods are applied. Statistical features are extracted from the breast thermograms and classified among Random Forest and Decision tree. Among the three methods, Adaptive Gaussian thresholding is giving a better accuracy of 94.67% and the other performance parameters such as precision, specificity and sensitivity are higher for the same as compared to the Decision Tree with the other two thresholding methods. For future work, automatic segmentation will be applied for obtaining more précised image which will improve the visual of the breast thermograms and assist in detecting the abnormality in the breast thermograms.

References

1. WHO, Cancer, URL: https://www.who.int/en/news-room/fact-sheets/ detail/cancer, 2021.
2. T.Z. Tan, C. Quek, G.S. Ng, E. Ng.: A novel cognitive interpretation of breast cancer thermography with complementary learning fuzzy neural memory structure, Expert Syst. Appl. 33 (3), 652–666 (2007).
3. O. D. Nurhayati.: Principal component analysis combined with first order statistical method for breast thermal images classification, 1 (2011).
4. L.S. Motta, A. Conci, R.C.F. Lima, E.M. Diniz.: Automatic segmentation on thermograms in order to aid diagnosis and 2D modelling, Proceedings of the Tenth Workshop em Informatica Medica, Belo Horizonte, pp. 1610-1619, MG, Brazil (2010).
5. Sharma, Ritam, Janki Ballabh Sharma, Ranjan Maheshwari, and Dumitru Baleanu.: Early anomaly prediction in breast thermogram by hybrid model consisting of superpixel segmentation, sparse feature descriptors and extreme learning machine classifier. Biomedical Signal Processing and Control 70, 103011 (2021).
6. Gogoi, Usha Rani, Mrinal Kanti Bhowmik, Anjan Kumar Ghosh, Debotosh Bhattacharjee, and Gautam Majumdar. Discriminative feature selection for breast abnormality detection and accurate classification of thermograms. In international conference on innovations in electronics, signal processing and communication (IESC), pp. 39-44. IEEE, 2017.
7. D. Sathish, S. Kamath, K. Prasad, R. Kadavigere;: Texture analysis of breast thermograms using neighbourhood grey tone difference matrix, International Journal of Bioinformatics Research and Applications 14 (1-2), 104–118 (2018).
8. Santana, Maíra Araújo de, Jessiane Mônica Silva Pereira, Fabrício Lucimar da Silva, Nigel Mendes de Lima, Felipe Nunes de Sousa, Guilherme Max Silva de Arruda, Rita de Cássia Fernandes de Lima, Washington Wagner Azevedo da Silva, and Wellington Pinheiro dos Santos.: Breast cancer diagnosis based on mammary thermography and extreme learning machines. Research on Biomedical Engineering 34, 45-53 (2018).
9. Sathish, Dayakshini, Surekha Kamath, Keerthana Prasad, and Rajagopal Kadavigere.: Role of normalization of breast thermogram images and automatic classification of breast cancer. The Visual Computer 35 (1), 57-70 (2019).

10. Karthiga, R., and K. Narasimhan.: Medical imaging technique using curvelet transform and machine learning for the automated diagnosis of breast cancer from thermal image. Pattern Analysis and Applications 24(3), 981-999 (2021).
11. A. Conci, Breast thermograms data, http://visual.ic.uff.br/en/ proeng/thiagoelias/ (2014).
12. Sahoo, Prasanna K., S. A. K. C. Soltani, and Andrew KC Wong.: A survey of thresholding techniques. Computer vision, graphics, and image processing 41(2), 233-260 (1998).
13. Filipczuk, Paweł, Marek Kowal, and Andrzej Obuchowicz.: Automatic breast cancer diagnosis based on k-means clustering and adaptive thresholding hybrid segmentation. In Image processing and communications challenges, vol. *3*, pp. 295-302. Springer, Berlin, Heidelberg (2011).
14. Otsu, Nobuyuki.: A threshold selection method from gray-level histograms." IEEE transactions on systems, man, and cybernetics 9(1), 62-66 (1979).
15. R. M. Haralick, K. Shanmugam, I. H. Dinstein.: Textural features for image classification, IEEE Transactions on systems, man, and cybernetics 6, 610-621 (1973).
16. X. Tang.: Texture information in run-length matrices, IEEE transactions on image processing 11 (7), 1602- 1609 (1998).
17. G. Thibault, B. Fertil, C. Navarro, S. Pereira, P. Cau, N. Levy, J. Sequeira, J.-L. Mari.: Shape and texture indexes application to cell nuclei classification, International Journal of Pattern Recognition and Artificial Intelligence 27(1), 1357002 (2013).
18. D. C. R. Novitasari, A. Lubab, A. Sawiji, A. H. Asyhar.: Application of feature extraction for breast cancer using one order statistic, glcm, glrlm, and gldm. Advances in Science, Technology and Engineering Systems Journal 4(4), 115-120 (2019).
19. Robnik-Sikonja, M., and I. Kononenko.: Theoretical and Empirical Analysis of ReliefF and RReliefF. Machine Learning 53, 23–69 (2009).
20. Chao, Cheng-Min, Ya-Wen Yu, Bor-Wen Cheng, and Yao-Lung Kuo.: Construction the model on the breast cancer survival analysis use support vector machine, logistic regression and decision tree. Journal of Medical System 38(10), 106–113 (2014).
21. Nguyen C, Wang Y, Nguyen HN.: Random Forest classifier combined with feature selection for breast cancer diagnosis and prognostic. J Biomed Sci Eng. 6(1), 551–560 (2013).