SCIENTIFIC PAPER

CrossMark

Singular value based characterization and analysis of thermal patches for early breast abnormality detection

Usha Rani Gogoi[1] · Mrinal Kanti Bhowmik[1] · Debotosh Bhattacharjee[2] · Anjan Kumar Ghosh[1]

Abstract

The purpose of this study is to develop a novel breast abnormality detection system by utilizing the potential of infrared breast thermography (IBT) in early breast abnormality detection. Since the temperature distributions are different in normal and abnormal thermograms and hot thermal patches are visible in abnormal thermograms, the abnormal thermograms possess more complex information than the normal thermograms. Here, the proposed method exploits the presence of hot thermal patches and vascular changes by using the power law transformation for pre-processing and singular value decomposition to characterize the thermal patches. The extracted singular values are found to be statistically significant ($p < 0.001$) in breast abnormality detection. The discriminability of the singular values is evaluated by using seven different classifiers incorporating tenfold cross-validations, where the thermograms of the Department of Biotechnology-Tripura University-Jadavpur University (DBT-TU-JU) and Database of Mastology Research (DMR) databases are used. In DMR database, the highest classification accuracy of 98.00% with the area under the ROC curve (AUC) of 0.9862 is achieved with the support vector machine using polynomial kernel. The same for the DBT-TU-JU database is 92.50% with AUC of 0.9680 using the same classifier. The comparison of the proposed method with the other reported methods concludes that the proposed method outperforms the other existing methods as well as other traditional feature sets used in IBT based breast abnormality detection. Moreover, by using Rank1 and Rank2 singular values, a breast abnormality grading (BAG) index has also been developed for grading the thermograms based on their degree of abnormality.

Keywords Breast cancer · Infrared breast thermography · Thermal patches · Singular value decomposition · Breast abnormality detection · Breast abnormality grading

Introduction

Early detection and diagnosis of breast diseases are crucial to reduce the incidence and mortality rate of breast cancer by improving the survival benefits. In India, the incidence rate of breast cancer increases in young age group and the cancer growth is very aggressive in women of younger age group [1]. Moreover, the vulnerability of young women towards the radiation risk and high false negative rate of the traditional gold standard method X-ray mammography necessitates exploring the efficiency of non-invasive and non-radiating infrared breast thermography (IBT) in early breast abnormality detection. Since breast cancer is asymptomatic until late in the disease process, regular breast health examination is necessary to identify any change in breast health. With the sensitivity of 90% [2], IBT can be used for routine examination of the breast health in asymptomatic patients for detecting the cases that require urgent medical attention. In literature [3, 4], it is reported that the thermograms which possess asymmetric temperature distributions signify the physiological dysfunction in patients' breasts most of the time [5]. Compared to these thermograms, the thermograms having increased nipple temperature, hot patches, and vascular changes may be more suspicious and indicate more severe breast problems [6, 7]. The appearance of these hot patches and vascular changes are due to the heat produced by the high metabolic activities of the blood vessels,

✉ Usha Rani Gogoi
ushagogoi.cse@gmail.com

[1] Department of Computer Science & Engineering, Tripura University (A Central University), Suryamaninagar, Tripura 799022, India

[2] Department of Computer Science & Engineering, Jadavpur University, Kolkata, West Bengal 700032, India

🖄 Springer

changes in blood perfusion, etc. in the suspicious regions [7, 8]. However, the intensity of heat radiation emitted by a breast mass may vary with its location within the breast [9], due to which a malignant tumor present deeper inside the breast may radiate poor radiation than a benign tumor present near the skin surface. Besides, the subtle temperature difference in breast thermograms may lead to an incorrect analysis of thermograms. So in order to identify the presence of breast abnormality in early stages, designing of an efficient computer-aided detection (CAD) system to analyze the thermal patterns may play a significant role.

As reported in the literature, the most common breast abnormality detection method is the bilateral asymmetry analysis of breast thermograms, which is based on the premise that the thermal patterns of the two breasts of an abnormal thermogram are noticeably asymmetric [10]. Although many contributions to the breast abnormality detection from infrared thermograms have been found in the literature [3, 11, 12], differentiating the abnormal thermograms from the normal one is still very challenging because of the subtle nature of temperature patterns in breast thermograms. The early and accurate detection of breast abnormality is important not only to provide a second opinion to the physicians in decision making but also for treatment planning.

In this study, we propose an automatic breast abnormality detection system to support the radiologists in decision making. Figure 1 depicts the flow of the proposed breast abnormality detection system. The proposed method is motivated by the fact that the abnormal thermograms have more high temperature regions or thermal patches in either breast than the normal thermograms, for which the abnormal thermograms have more spatial information than the normal thermograms. To prove this fact, "Association of singular values with image complexity of breast thermograms" section performs the complexity analysis of the processed breast thermograms which concludes that the abnormal thermograms possess more complex information than the normal thermograms. The proposed system aims at discriminating the abnormal thermograms with suspected abnormality and also to predict the degree of abnormality in these thermograms. Such an automatic system can help to reduce the radiologists' burden and examination time with the additional advantages of objectivity and diagnosis quality. Moreover, it can help to rapidly identify the cases that require urgent medical attention and further evaluation. The key contributions of this paper are summarized as follows.

1. We propose a simple yet effective breast abnormality detection method. Unlike the conventional method of using texture feature based bilateral asymmetry analysis, the proposed method makes use of the candidate thermal patches to signify the breast abnormality. A set of singular values (SVs) has been extracted from the preprocessed images to quantify the abnormality present in each thermogram.
2. The classification performance of the proposed breast abnormality detection system has been evaluated with both in-house and public breast thermogram databases namely Department of Biotechnology-Tripura University-Jadavpur University (DBT-TU-JU) [13] and Database of Mastology Research (DMR) [14] by using a series of seven state-of-the-art classification systems.
3. The potential of the SVs in breast abnormality detection is compared with some other feature sets and the evaluation results reflect that the proposed method outperforms the existing methods.
4. A breast abnormality grading (BAG) index has been designed by using 1st and 2nd rank SVs to grade the abnormal thermograms into mild abnormal and severely abnormal based on the degree of abnormality present in it.

The rest of the paper is organized as follows. A related work on IBT based breast abnormality detection is given in "Related work" section. The detail description of the proposed breast abnormality detection method is provided in "Proposed method" section. "Experimental results" section provides the details of the experimental databases and the experimental results. The comparative studies are presented in "Discussion" section. Finally, "Conclusion" section concludes the paper.

Related work

IBT with the characteristics of showing the abnormal thermal patterns or hot thermal patches as the suspicious breast regions draw the attention of the researchers to
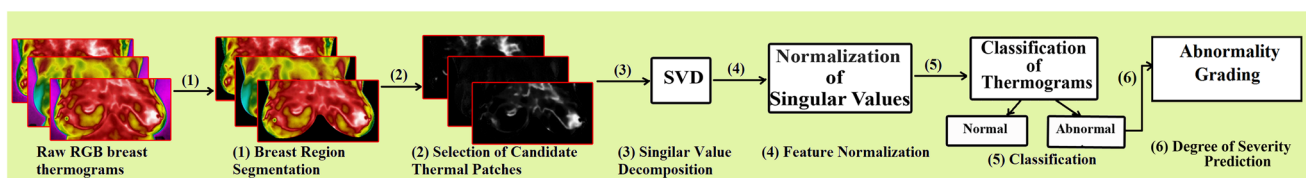

Fig. 1 The overview of the proposed breast abnormality detection system

use it in breast abnormality detection. Bilateral asymmetry analysis is the most widely investigated and classical method of breast abnormality detection from thermograms. In literature [3, 11], there are several prior investigations in asymmetry analysis based breast abnormality detection from thermograms. A summary of the existing works on asymmetry analysis based breast abnormality detection has been illustrated in Table 1. Table 1 summarizes the system protocol like the feature sets, the size of

the experimental dataset, the classifier algorithm and the performance scores of the existing methods [5, 15–27].

Even though a lot of work had been done in bilateral asymmetry based breast abnormality detection, this conventional method of bilateral asymmetry analysis cannot detect the presence of breast abnormalities in breast thermograms where both breasts are suffering from abnormalities. It is because, in thermograms where both breasts are suffering from breast abnormalities, the features that measure the bilateral asymmetry possess a very minute difference, which

Table 1 Summary of existing bilateral asymmetry analysis based breast abnormality detection methods from thermograms

Authors	Method used	Dataset	Classifier used	Performance measures		
				Acc.	Sens.	Spec.
Schaefer et al. [5]	Statistical and texture features	146 (29M 117B)	Hybrid Fuzzy rule based classifier	80.00%	–	–
Ng et al. [15]	Temperature data and patient's clinical biodata	82 (30N, 48B, 4C)	RBFN	80.95%	81.20%	88.20%
Mookiah et al. [16]	Texture features and DWT	50 (25N, 25M)	DT, FS	93.30% 93.30%	86.70% 86.70%	100% 100%
Acharya et al. [17]	Texture features	50 (25N, 25C)	SVM	88.10%	85.71%	90.48%
Francis et al. [18]	Texture features	27 (9N, 18C)	FNN	85.19%	88.89%	77.78%
Francis et al. [19]	Statistical and texture features	36 (24N, 12M)	SVM	83.30%	83.30%	83.30%
Francis et al. [20]	Statistical and texture features	22 (11N, 11C)	SVM	90.91%	81.82%	100%
Araujo et al. [21]	Temperature based interval symbolic features	50 (14M, 19B, 17Cy)	Distance based classifier	–	85.70%	86.50%
Garduno-Ramon et al. [22]	Temperature features	454 (414N, 40Ab)	–	–	86.84%	89.43%
Gaber et al. [23]	Statistical, Gabor, texture features	63 (29N, 34M)	SVM	Average Acc. of four scenarios: 88.41%		
Zadeh et al. [24]	Patient's age, mean, variance, kurtosis, skewness, entropy, difference between two breasts and thermal pattern of the breasts	200	ANN	70.00%	50.00%	75.00%
Sathish et al. [25]	Statistical and texture features	80 (40N, 40Ab)	SVM	90.00%	87.50%	92.50%
Borchartt et al. [26]	Range temperature, mean temperature, Standard deviation, quantization of higher tone	28	SVM	85.71%	95.83%	25.00%
Lashkari et al. [27]	23 features including statistical, morphological, frequency domain, histogram and GLCM based features	67	AdaBoost	85.33% (L) 87.42% (R)	63.33 (L) 68.56 (R)	90.83 (L) 92.14 (R)

Acc. accuracy, *Sens*. sensitivity, *Spec*. specificity, *N* normal, *B* benign, *C* cancer, *M* malignant, *Cy* cyst, *L* left, *R* right, *Ab* abnormal, *DWT* discrete wavelet transform, *RBFN* radial basis function network, *DT* decision tree, *FS* fuzzy Sugeno, *FNN* feed forward network

ends up concluding the absence of abnormality. Hence, in contrast to bilateral asymmetry analysis for abnormality detection, our proposed method has used the SVs to characterize these thermal patterns irrespective of their location in either breast. Thus, the proposed method may improve the accuracy of breast abnormality detection by increasing the number of true positives or by detecting the abnormalities in those breast thermograms, where bilateral asymmetry analysis fails to detect.

Proposed method

Given a breast thermogram, a set of efficient features is required to be extracted for detection of abnormality. Considering that our work focuses on the generation of efficient features from breast thermograms that can provide better performance over the reported works. The proposed breast abnormality detection system primarily consists of four sub-steps. The first sub-step involves the pre-processing and the segmentation of the breast regions from raw breast thermograms. The second sub-step does the identification of the candidate thermal patterns from the segmented images by doing hue–saturation–value (HSV) transformation and power law transformation of the images. The third sub-step decomposes the preprocessed images into SVs to quantitatively characterize the candidate thermal patterns of the breast thermograms. In fourth sub-step, the extracted SVs are fed to seven different classification methods to find the most efficient classifier for our proposed method in differentiating the abnormal thermograms from the normal ones.

Furthermore, based on the 1st and 2nd rank SVs extracted from each breast thermogram, a BAG index has also been developed to predict the degree of abnormality it possesses. Detail description of each sub-step is provided in the following sub-sections.

Pre-processing and segmentation of breast region

Pre-processing of breast thermograms

In designing a computer aided breast abnormality detection system, the pre-processing of the raw thermograms is very crucial since the raw breast thermograms generally contain some irrelevant parts like neck portion, area underneath the breasts and the background, etc. Hence to make the raw thermograms in "Rainbow HC" color pallet more suitable for further processing and to enhance the performance of the proposed system, the breast thermograms are manually cropped to discard the irrelevant regions. For the subsequent steps, all the breast thermograms are resized to have a resolution of 200×400 pixels. The pre-processing of the breast

thermograms has been followed by the segmentation of the breast regions.

Segmentation of breast region

The primary objective of segmenting the breast region from a breast thermogram is to discard all those portions that are not belonging to the breasts so that the prediction accuracy of the breast abnormality detection system gets improved. Infrared breast thermograms are poor in contrast for which it lacks clear edges and it is amorphous in nature [28], which makes automatic segmentation of breast region very tedious. Although a lot of work have been done for automatically segmenting the breast region from a breast thermogram, due to the unclear lower parabolic boundaries in most of the cases present in our dataset, these methods are not found to be robust. Therefore by considering the fact that the lower breast boundaries and the inframammary folds should not be getting lost in segmentation of breast thermograms, we have adopted a semi-automatic segmentation method. This semi-automatic segmentation method requires the human interaction for selecting the lower parabolic boundaries of both breasts. Algorithm 1 summarizes the procedure of segmenting the breast region from a breast thermogram. Figure 2 shows the segmented breast regions of some sample breast thermograms.

Algorithm 1: The Breast Region Segmentation Algorithm

Input: RGB image I

Output: Segmented Breast Region S

Step 1. $I' \leftarrow$ Compute the blue difference Chroma component Cb by using the RGB to YCbCr Image conversion formulae described in [29, 30]

Step 2. $B \leftarrow$ Convert I' to binary image to distinguish the body area from the background by using thresholding segmentation technique

Step 3. $P \leftarrow$ Select the lower breast boundary points manually on I' where the coordinates of the last point must be the first point so that it creates a polygon

Step 4. $X \leftarrow$ Generate a binary mask of the polygon, P

Step 5. $M \leftarrow B - X$, M is the binary breast mask

Step 6. $S \leftarrow I * M$

Identification of candidate thermal patches

The goal of this step is to enhance the segmented breast thermograms to make the candidate thermal patches more prominent by neglecting the unnecessary surrounding regions. The candidate thermal patches are those thermal patterns that are characteristically in contrast to the surrounding regions on its color and intensity that can be used as an important fact for separating these candidate patterns from the surrounding areas. As illustrated in Fig. 3a–c, the temperature radiations emitted by different regions of a breast are represented with

Fig. 2 **a** Manually selected lower parabolic curve, **b** generated breast mask, **c** segmented breast region
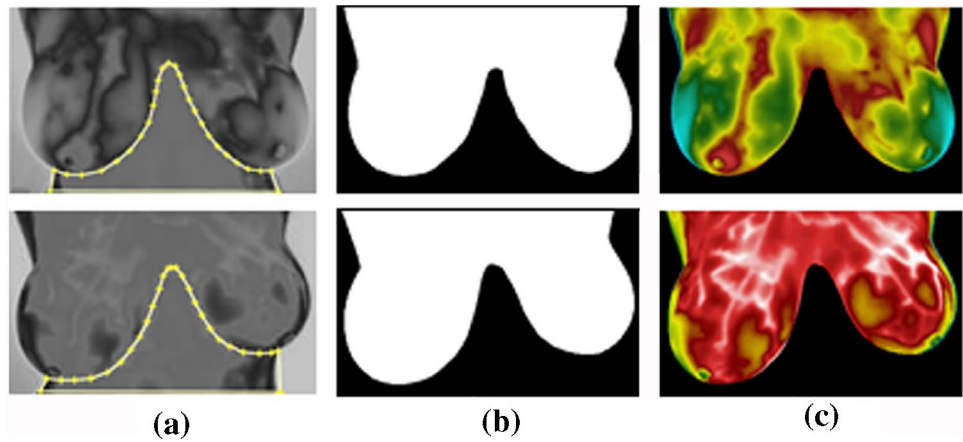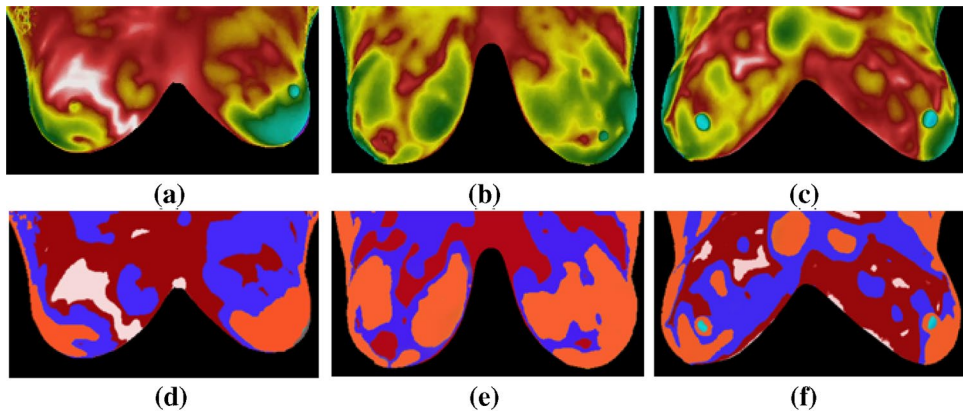

(a) (b) (c)

Fig. 3 **a–c** Some sample breast thermogram images; **d–f** corresponding labeled images with candidate thermal patches

(a) (b) (c)

(d) (e) (f)

different pseudo colors. For better description of the candidate thermal patterns in thermograms, we have labeled each temperature region with different colors as shown in Fig. 3d–f. Among all the pseudo colors, the white region corresponds to the highest temperature region (hotspot) indicating the anomalous region of a breast and this highest temperature region along with its surrounding reddish white regions are considered as the candidate thermal patches. Figure 3a, d shows a breast thermogram with a malignant tumor in the right breast. Figure 3b, e shows the breast thermogram of a healthy person with no candidate thermal patches and Fig. 3c, f shows the breast thermogram of a healthy person with some candidate thermal patches. In order to make these candidate thermal patterns more prominent, two methods have been used here.

RGB to HSV transformation

In a breast thermogram image, the color component does not carry important information, for which these color information can be removed to reduce the processing complexity of breast thermograms. Unlike the RGB color model, the HSV model has the advantages in that the intensity is decoupled from the color information, for which object description in terms of

these components is easy. Besides, the hue and saturation are closely related to the way in which human observes color [30]. Figure 4 illustrates the histogram of each channel of an RGB and HSV color spaced breast thermogram. As depicted in Fig. 4a, it can be stated that in the RGB breast thermogram, all three color components are dominant over the entire dynamic range; no channel is efficient in highlighting the highest temperature regions. On the contrary, Fig. 4b illustrates that if converted to HSV color space, each of the three color components is dominant in different ranges. Hue is found to be more or less dominant over the entire dynamic range, while saturation is dominant over the higher scale and the value is found to be dominant over the middle to higher scale of the dynamic range. Based on the distribution range of each channel in the HSV breast thermogram, we have found that saturation channel will be more suitable for our research work to characterize the higher temperature regions. Hence, for developing a human visual system based breast abnormality detection method, the saturation channel (S) has been extracted from each segmented breast thermogram by using the Eq. (1) [30, 31],

$$S = \begin{cases} 1 - \frac{3}{R+G+B}\min(R, G, B), & if\ R+G+B > 0, \\ 0 & if\ R+G+B = 0. \end{cases}$$
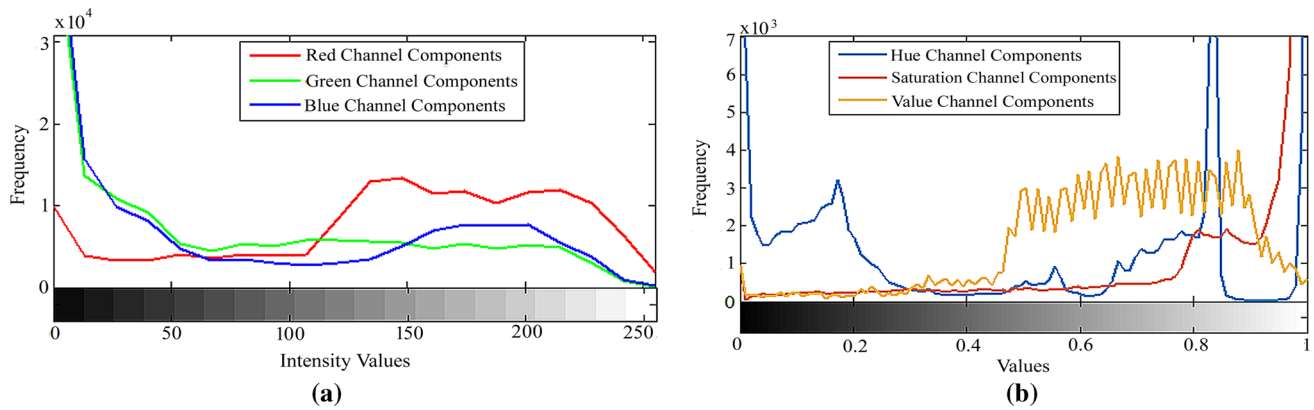(1)

Fig. 4 The histogram plot of each channel of a breast thermogram image in **a** RGB color space and in **b** HSV color space

The saturation refers to the pureness of the color or the amount of whiteness in the color mixture [30] and hence, the higher temperature regions having the maximum intensity values are poorly saturated. The hue, saturation and value components obtained from a breast thermogram image are illustrated in Fig. 5. As shown in Fig. 5b, it has been seen that the candidate thermal patches have the highest contrast in the saturation channel and these patches appear as a group of pixels with local minimal intensity surrounded by higher intensity pixels. However, in contrary to the fact that the hot patch bears the highest temperature or intensity value, saturation channel makes the hot patches to have the lowest intensity values due to the poor saturation. Hence, the saturation channels of all breast thermograms are complemented so that the higher temperature regions are represented with the higher intensity values as shown in Fig. 5d and it is denoted as I_s. Here it is worth mentioning that only the pixel values inside the breast region of the saturation channel get complemented for which the non-breast region (the black region of the breast mask shown in Fig. 2b) remain same before and after performing the complementation.

Power law transformation

In thermal images, due to the absence of sharp transition of intensity values from one region to another, the thermal patterns do not have sharp boundaries and it causes the appearance of some unnecessary regions in I_s as shown in Fig. 5d. Hence to further improve the representation power of the channel and to make the candidate thermal patches more prominent by removing all the unnecessary portions, the power law transformation has been performed. The power law transformation reduces the insignificant spread of the patches in I_s to make the candidate patches more distinguishable. The power law transformed image, I_p is obtained by,

$$I_P(i,j) = c \times I_s(i,j)^\gamma, \tag{2}$$

where c, γ are positive constants.

The power law transform changes the dynamic range of I_s. It is worth mentioning that for any γ > 1, power law transform increases the bandwidth of the high intensity values at the cost of the low pixel values and for γ < 1, it enhances the low intensity value while decreasing the bandwidth of the
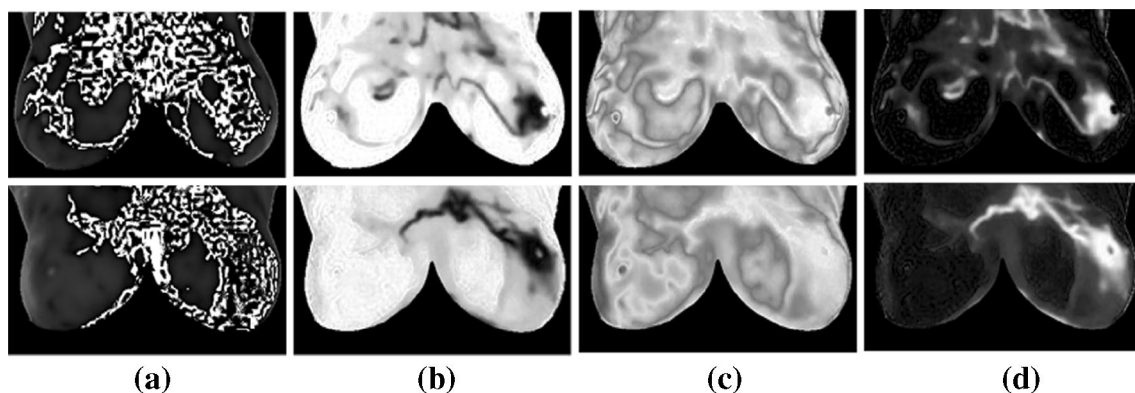

Fig. 5 **a** Hue channel, **b** saturation channel, **c** value channel and **d** complemented saturation channel of a breast thermogram

high intensity values. Since our motive is to suppress the low intensity pixel values by enhancing the high intensity pixels, we have used positive values of γ. For our experimental purpose, the γ value is set to 3.5. Moreover, the value of c is fixed at 1 as for any value of c not equal to 1, scaling significantly affects the dynamic range of the pixel values of I_s.

Feature extraction

Feature extraction is probably the single most essential step in achieving high accuracy in breast abnormality prediction. In pattern recognition, it is desirable to extract features that focus on discriminating between classes. In this section, our main objective is to extract an efficient feature set from the candidate thermal patches that can significantly characterize the breast abnormality with limited number of features. To deal with this problem, we have utilized the image compression property of singular value decomposition (SVD), where it can represent an image with a limited number of SVs that can preserve the useful information of I_p. The extraction of SVs, their significance, their association with image complexity and normalization are detailed in subsections.

Singular value decomposition

SVD is a linear transformation of an $M \times N$ matrix, which refactors the matrix into three component matrices. Thus, SVD decompose the I_p into three component matrices, *U, S* and *V* such that

$$I_p = USV^*,\tag{3}$$

where *U* is an M x M orthogonal matrix containing the left singular vectors of I_p, *V* is an $N \times N$ orthogonal matrix containing the right singular vectors of I_p, *S* is an $M \times N$ diagonal matrix, in which the nonnegative entries along the diagonal of *S* are the singular values of I_p. The singular values $\sigma_1, \sigma_2 \dots \sigma_k$, $k = \min(M, N)$ are unique. If *r* is the rank of the matrix I_p, then,

(i) $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$,

(ii) $\sigma_{r+1} = \sigma_{r+2} = ... = \sigma_n = 0$.

Proposition *The low rank SVs of the abnormal thermograms are larger than the low rank SVs of the normal thermograms.*

Since the presence of thermal patches signifies the presence of breast abnormality, let us assume the abnormal thermograms bear more candidate thermal patches than the normal thermograms. Let, the normal thermogram be *I* and abnormal thermograms be $I + \Omega$, where Ω: additional candidate thermal patches.

Now, the properties of 2-norm of a matrix can be used. However, in practice it is difficult to compute the 2-norm of a matrix due to the unavailability of direct formula. Hence instead of directly computing the 2-norm of a matrix, we use the concepts of vector norm to compute the 2-norm of a matrix. For doing this, we need to treat the $M \times N$ elements of any matrix *A* as the elements of an *MN* dimensional vector, then compute the 2-norm of the vector as follows [32],

$$\|A\|_2 = \|A\|_F = \left\{ \sum_{i=1}^{M} \sum_{j=1}^{N} \left| a_{ij} \right|^2 \right\}^{\frac{1}{2}}.\tag{4}$$

For easiness, we denote *I* as *A* and $I + \Omega$ as *B*. If *A* and *B* are of same order and Ω: additional candidate thermal patches in *B*, then,

$$b_{ij} > a_{ij}, \quad a_{ij} \in A, b_{ij} \in B, \quad i,j \in \Omega \Rightarrow \sum_{i=1}^{M} \sum_{j=1}^{N} \left| b_{ij} \right| > \sum_{i=1}^{M} \sum_{j=1}^{N} \left| a_{ij} \right|$$

$$\Rightarrow \left\{ \sum_{i=1}^{M} \sum_{j=1}^{N} |b_{ij}|^2 \right\}^{\frac{1}{2}} > \left\{ \sum_{i=1}^{M} \sum_{j=1}^{N} |a_{ij}|^2 \right\}^{\frac{1}{2}} \Rightarrow ||B||_2 > ||A||_2.$$

Thus,

$$\|I + \Omega\|_2 > \|I\|_2.\tag{5}$$

Another definition to compute 2-norm of a matrix is as follows [33, 34],

$$||A||_2 = \sqrt{\text{Maximum eigenvalue of } A'A}$$
$$= \text{Largest singular value of matrix A}.\tag{6}$$

Thus, Eqs. (5) and (6) imply that,

$$||I + \Omega||_2 > ||I||_2$$

$\Rightarrow$ Largest singular value of $I + \Omega$ > Largest singular value of *I*.

If we denote largest SV or Rank1 SV of a matrix as $\underline{\sigma}$, then

$$\underline{\sigma}(I + \Omega) > \underline{\sigma}(I).\tag{7}$$

Based on the Eq. (7), we can conclude that the first SV of an abnormal thermogram is larger than the first SV of a normal thermogram. In the same way, we can show that other low rank SVs of an abnormal thermogram are also larger than the corresponding lower rank SVs of the normal thermograms. Here, we mention the term low rank because, as rank increases the magnitude of the SVs tends to be zero.

The first 20 (from Rank1 to Rank20) SVs extracted from all breast thermograms of the normal and abnormal groups are plotted in Fig. 6. As illustrated in Fig. 6a, it has been seen that the magnitudes of the rank *r* SVs of the abnormal thermograms are larger than the magnitudes of the

corresponding *r* rank SVs of the normal thermograms. However as the rank increases, the discriminating power of SVs gets reduced. In Fig. 6b, the rank wise average of the SVs obtained from the thermograms of the abnormal group is plotted against the rank-wise average of the SVs of the thermograms of normal group and it also depicts the discriminating power of lower rank SVs.

Statistical significance of rank wise singular values

Extraction of SVs from breast thermograms is followed by the testing of the statistical significance of these SVs in distinguishing breast thermograms. For this, 2-sample *t* test with significance level 0.1% and null hypothesis, Ho: the abnormal and normal thermograms have equal means has been used. Due to the limitation of space to show the statistical significance of all the extracted SVs, the statistical significance of first 10 SVs i.e., from Rank1 to Rank10 for both normal and abnormal thermograms have been demonstrated in Table 2. Table 2 illustrates that the average magnitude of the SVs (in each rank) of the abnormal thermograms is significantly

greater than the average magnitude of the SVs of the corresponding rank of the normal thermograms. Moreover as shown in Table 2, the SVs of all ranks, obtained from all the normal and abnormal thermograms are statistically significant with P-value < 0.001 in breast abnormality detection.

Association of singular values with image complexity of breast thermograms

This section analyzes the association of the SVs with the degree of image complexity that a breast thermogram bears. In the absence of any standard definition of image complexity, here the complexity of a breast thermogram is considered as the amount of information present after preprocessing the thermograms. Hence, the image complexity is not measured from the raw breast thermograms; instead the thermograms possessing the candidate thermal patches are used for measuring the image complexity. Three different measures are used here to compute the complexity of processed breast thermograms. The first measure is based on the Shannon's definition of information, known as entropy [35].
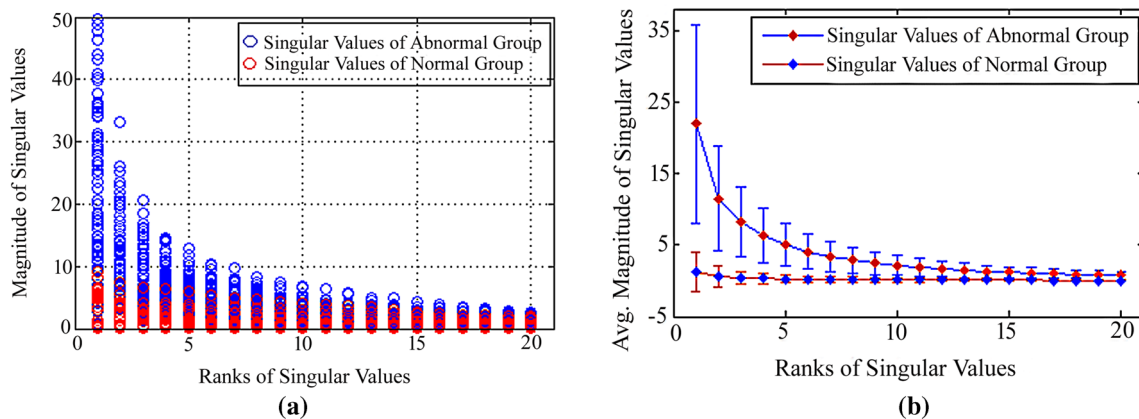

Fig. 6 **a** Rank-wise magnitudes of SVs of all normal and abnormal breast thermograms of DBT-TU-JU dataset, **b** Rank-wise average of the SVs of normal and abnormal breast thermograms of DBT-TU-JU dataset

Table 2 Average of SVs in each rank

Rank	DBT-TU-JU			DMR		
	Normal group	Abnormal group	P-value	Normal group	Abnormal group	P-value
R1	3.06 ± 2.66	23.49 ± 12.37	3.1e−21	1.94 ± 1.76	18.33 ± 10.33	4.5e−33
R2	1.94 ± 1.78	12.22 ± 6.65	1.8e−21	1.38 ± 1.10	9.471 ± 5.121	6.3e−33
R3	1.43 ± 1.33	8.846 ± 4.479	6.7e−19	1.10 ± 0.85	6.751 ± 3.809	5.9e−30
R4	1.22 ± 1.21	6.919 ± 3.475	2.5e−20	0.94 ± 0.73	5.133 ± 2.863	9.0e−29
R5	1.07 ± 1.11	5.465 ± 2.800	6.4e−19	0.87 ± 0.66	4.222 ± 2.557	3.5e−25
R6	0.93 ± 0.97	4.535 ± 2.330	1.9e−18	0.79 ± 0.60	3.585 ± 2.197	5.6e−24
R7	0.86 ± 0.91	3.863 ± 2.092	1.3e−16	0.74 ± 0.56	3.161 ± 2.025	4.3e−22
R8	0.79 ± 0.86	3.260 ± 1.821	3.8e−15	0.69 ± 0.53	2.746 ± 1.856	6.9e−20
R9	0.73 ± 0.79	2.815 ± 1.596	3.1e−14	0.65 ± 0.49	2.454 ± 1.681	5.9e−19
R10	0.67 ± 0.74	2.472 ± 1.465	4.9e−13	0.61 ± 0.47	2.184 ± 1.527	1.2e−17

Let $p(i) = n(i)/N$ be the probability of gray level i, where $n(i)$ is the number of pixels having gray level i and N is the total number of pixels in a breast thermogram. Then, Shannon's entropy can be defined as,

$$H = -\sum_{i=0}^{255} p(i) \times \log p(i), \quad \forall i \in I_p. \tag{8}$$

The second measure is the spatial information (SI) [36], which is an indicator of edge energy and it is computed as follows.

Let H_s and V_s denote the edge images generated by applying the horizontal and vertical kernels of Sobel, then the spatial information SI_{mean} of a breast thermogram is given by,

$$SI_{mean} = \frac{1}{N} \sum S_i, \quad \forall i \in I_p, \tag{9}$$

where N is the total number of pixels in I_p and

$$S_i = \sqrt{H_s^2 + V_s^2}, \quad \forall i \in I_p. \tag{10}$$

The third measure is the fractal dimension (FD) [37], which observes how the number of boxes deviates as the grid becomes finer by applying a box-counting algorithm. To determine the FD of a pattern, it is imagined that the pattern is laid on an equally spaced grid, and then the number of boxes required to cover the whole set is counted [38]. Mathematically, it is computed as follows.

Let B is the number of boxes that covers the candidate thermal patches and r is the magnification or the inverse of box size, then FD is the slope of the line when the value of $log(B)$ is plotted on Y-axis against the value of $log(r)$ on the X-axis and it is given by,

$$FD = \frac{\log(B)}{\log(r)}. \tag{11}$$

After extracting the image complexity values of the thermograms of both normal and abnormal group, it has been seen that the abnormal thermograms possess significantly higher entropy (ent), SI and FD values than that of the normal breast thermograms, i.e., based on these three complexity measures, we can conclude that the abnormal thermograms are more complex than the normal thermograms. The mean and standard deviation of the values of these complexity metrics for the thermograms of both normal and abnormal group are demonstrated in Table 3. Moreover, as illustrated in Table 3, all the complexity metrics are statistically significant (p < 0.001, t test) in showing that the abnormal thermograms are having higher image complexity than the normal thermograms.

Now, the association of the computed SVs with the complexity of the thermograms has been evaluated by using the Pearson Correlation measure [39], which is a statistical tool used to measure the degree to which the variables are associated with each other. The correlation between the rank r SVs of all thermograms and their corresponding image complexity measures are analyzed and it has been found that the SV of any rank is positively correlated with the image complexity that means for an image with higher complexity, the SVs will also be greater. Thus, the abnormal breast thermograms have higher SVs than that of the normal thermograms. The correlation of the Rank1 SVs of all the normal and abnormal thermograms with the corresponding image entropy, SI and FD has been illustrated in Fig. 7 for both DBT-TU-JU and DMR databases.

Normalization of singular values

As demonstrated in Table 2, the SVs of different ranks exhibit significant variation in their range, which necessitates the normalization or rescaling of the SVs before feeding them to various classifiers. Moreover, the normalization of all SVs to a fixed range ensures that each feature contributes proportionately to the final match score.

Let f and f' denote a feature vector before and after normalization. The normalized feature vector f' is computed through,

$$f_i'{}_{i \in N} = \frac{f_i - \min(f)}{\max(f) - \min(f)}, \tag{12}$$

where N is the number of thermograms.

Classification of breast thermograms

In a CAD system, the feature extraction is followed by the classification of the images based on the extracted features. Although all the SVs are found to be statistically significant, utilization of all the SVs may degrade the accuracy of the

Table 3 Average values of image complexity metrics

| Dataset used	Complexity measures	Abnormal group	Healthy group	P-value
DBT-TU-JU	Entropy	2.4977±0.833	0.7752±0.607	2.9052e−23
	Spatial information	0.0258±0.008	0.0120±0.005	3.1442e−18
	Fractal dimension	0.9121±0.165	0.4838±0.207	2.0367e−16
DMR	Entropy	1.1359±0.526	0.2533±0.247	1.9958e−28
	Spatial information	0.0151±0.005	0.0085±0.004	1.0370e−14
	Fractal dimension	0.8661±0.141	0.5209±0.142	4.2832e−17
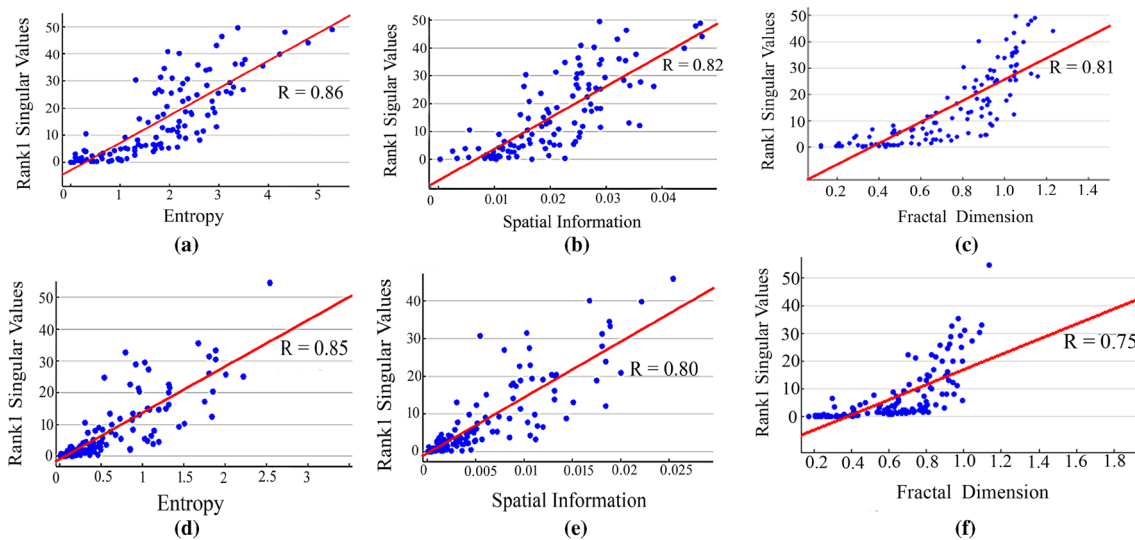
Fig. 7 **a–c** Correlation of Rank1 SVs of all breast thermograms of DBT-TU-JU dataset with the image complexity measures, **d–f** correlation of Rank1 SVs of all breast thermograms of DMR dataset with the image complexity measures

system as the discriminating power of the SVs gets reduced with increasing ranks. Hence, to obtain a right combination of SVs and classifier to attain best prediction performance, different combination of SVs with seven most widely used classifiers including support vector machine (SVM) with three different kernels [radial basis function (RBF), polynomial, linear], K-nearest neighborhood (KNN), decision tree (DT), artificial neural network (ANN), random forests (RF), AdaBoost (AB) and linear discriminant analysis (LDA) have been investigated.

Designing of breast abnormality grading (BAG) index

Determination of degree of abnormality in an abnormal thermogram is very much crucial for early detection and diagnosis of the severely abnormal cases. Considering this, the classification of thermograms as normal and abnormal is followed by the designing of a BAG index. For designing the BAG index, among all the SVs of a thermogram, we consider the first two SVs, SV1 and SV2, which are found to be highly discriminative as shown in Table 2. By using these two SVs: SV1 and SV2, the BAG index is defined as-

$$BAG\ index = SV_1 + SV_2. \tag{13}$$

As illustrated in "Association of singular values with image complexity of breast thermograms" section, the SVs are highly correlated with the image complexity for which the SVs SV1 and SV2 are also highly correlated with abnormality. Hence, thermograms with less abnormality will have smaller values of BAG index compared to the BAG index values of the thermograms having severe abnormality. Thus,

based on the values of BAG index, we can predict the degree of abnormality in a breast thermogram.

Experimental results

Experimental databases

Both publically available and in-house acquired databases were used to evaluate the performance of our proposed breast abnormality detection method. Two breast thermogram databases namely DBT-TU-JU [13] and DMR [14] have been used. The DBT-TU-JU database is our own developed database in collaboration with Regional Cancer Centre (RCC), Agartala Government Medical College (AGMC), Tripura, India. To acquire the thermograms, a standard acquisition protocol suite had been designed [13] that comprises of a number of important parameters: patient preparation, patient acclimation, patient intake form, examination room condition, patient position and acquisition views. The breast thermograms were acquired by using the FLIR T650sc thermal camera with thermal sensitivity of <20 mK @ 30 °C and image resolution of 640×480 pixels. Currently, the database comprises the thermograms of both healthy and pathological subjects in the age group of 21–80 years. Moreover, the database is also annotated with the ground truth images of the suspicious hot regions. From this database, a dataset of 120 breast thermograms (70 abnormal and 50 normal) in frontal view has been considered for the experimental purpose. The set of abnormal thermograms comprises of 35 thermograms having benign tumors, 15 thermograms having malignant tumors and

20 other thermograms suffering from other breast problems like breast pain, feeling of solid mass, bloody discharge, pus formation etc.

For validating the performance of the designed BAG index in breast severity prediction, the abnormal thermograms are clinically categorized into mild abnormal (MA) and severely abnormal (SA). Due to the incapability of X-ray mammography in breast abnormality detection in all suspected cases of DBT-TU-JU database, the categorization of breast thermograms has been done based on patient history, clinical findings, X-ray mammography report and Fine Needle Aspiration Cytology (FNAC) report. Considering the findings of these modalities, all the abnormal thermograms are categorized into MA and SA. The abnormal thermograms whose mammography or FNAC reports show the presence of either benign or malignant tumor are labeled as SA. On the other hand, the abnormal thermograms whose FNAC reports are not available and mammography could not reveal the presence of any tumor or calcifications, but clinically they are found to be abnormal or having some disease related symptoms like feeling of solid and lumpy structures, experiencing pain for a long period of time are labeled as MA. Moreover, the abnormal thermograms of those patients whose mammography reports are normal, but they are experiencing blood or pus discharge for very long periods of time are also labeled as SA.

The publically available DMR database [14] contains the breast thermograms of 287 subjects. The breast thermograms of almost 47 subjects are labeled as 'Sick' and remaining are labeled as 'Healthy'. It is reported that the breast thermograms were captured using FLIR SC-620 Infrared camera with temperature sensitivity < 0.04 °C and image resolution: 640×480 pixels. Here for experimental purpose, a dataset of 100 normal (healthy) thermograms and 45 abnormal (sick) thermograms has been used.

Evaluation metrics

To facilitate the performance evaluation of our proposed breast abnormality detection system, three evaluation metrics: accuracy, sensitivity and specificity have been used that establish the superiority of our proposed method. These evaluation parameters are given as follows,

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{14}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{15}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{16}$$

where TP, TN, FP and FN indicate true positive, true negative, false positive and false negative respectively. The

sensitivity is the proportion of positive (abnormal) cases that are correctly identified as positive, specificity is the proportion of negative (normal) cases that are correctly identified as negative and the accuracy is the proportion of total number of cases that are correctly classified. Thus, if both sensitivity and specificity are high (low), the accuracy will also be high (low). But, if any of the sensitivity or specificity is high, then the accuracy will be biased to any one of these measures.

Assessment of classification performance

The classification performance of the proposed breast abnormality detection system has been evaluated with seven 'state-of-the-art' classification systems including: SVM, DT, KNN, ANN, RF, LDA and AB employing tenfold cross-validation. A brief description on designing of each of these classifiers except LDA is provided below. In LDA, no parameter tuning is done.

Support vector machine

For the SVM classifier, the performance has been tested with three different kernels: RBF, polynomial and linear kernel. Except for the linear kernel, the parameters of both polynomial and the RBF kernel are altered to obtain better classification accuracy. Among different values of sigma, the SVM with RBF kernel shows the highest accuracy against sigma = 2. Similarly, SVM with polynomial kernel provides the best accuracy against the polynomial order = 3.

Decision tree

In DT, except 'MaxNumSplits' parameter, other parameters are set as default to obtain the optimal classifier to better fit the data. For tuning DT, the maximum number of splits is set with a number ranging from 1 to 15 and found that with higher number of splits, our model can perfectly predict the train data, but fails to predict the test data. However, the maximum accuracy with train and test data is obtained with split number 5.

K-nearest neighborhood

To obtain the optimal number of neighborhood K for the KNN classifier, the classification accuracy of the proposed system with K = 2 to K = 10 has been evaluated with cross-validation and we have found that the highest classification accuracy was obtained with K = 9. Thus, KNN classifier with K = 9 is used as the optimal model.

Artificial neural network

For implementing ANN, we used pattern recognition network which is a feed-forward network, where the size of the input, output and output layers are initially set to zero. During training, according to training data, sizes of these parameters are adjusted. The number of hidden layers is varied from 1 to 10 and best prediction performance with the testing set is obtained with ten hidden layers. The designed network used 'Scaled Conjugate Gradient' training function to train the network.

Random forest

RF is an ensemble tool that combines multiple decision trees to get a more accurate and stable prediction. Here, *square root(total number of features)* is used as the maximum number of features in each decision tree. For obtaining the better prediction performance, different number of decision trees are used to model the RF classifier and the maximum classification was obtained with 20 numbers of decision trees.

AdaBoost

AB constructs a robust classifier by iteratively adding multiple weak classifiers. We used AdaBoost M1 classifier for binary classification. In each round of training, the classification performance of the ensemble is enhanced by adding a new weak learner. Here, one level decision tree also known as decision stump is used as weak learner to create the ensemble and the ensemble undergoes 100 learning iterations to achieve the optimal performance.

As aforementioned, the classification accuracy of the proposed system has been evaluated with different numbers of SVs so that we can make a choice of optimal number of SVs to obtain the best classification performance. However, instead of using all SVs, the performance of the proposed system has been evaluated with maximum of 30 SVs, since beyond this the performance of the system gets degraded noticeably. The maximum of the evaluation metric values obtained with each classifier for different number of SVs over 20 iterations is demonstrated in Table 4. The set of observations made from Table 4 are as follows,

1. Different classifiers show their best performances either with 1, 2 or 5 SVs. However in most of the classifiers, the highest prediction performance has been obtained with two SVs in both the DBT-TU-JU and DMR datasets. The performances of all the classifiers either get degraded or remain same when the number of SVs gets increased.

2. If the classification performance of a classifier is same for both 1 and 2 SVs, then their value of sensitivity is used to break the tie and the classifier with the higher sensitivity is considered to be better.

3. In DBT-TU-JU dataset, SVM with Polynomial kernel and ANN give the highest prediction accuracy of 92.50% with two SVs. Moreover in this dataset, the proposed system has obtained the prediction accuracy more than or equal to 90% with all other classifiers.

4. In DMR dataset also, the highest prediction accuracy of 98.00% has been obtained with SVM using polynomial kernel and ANN with two SVs. Except for LDA, all other classifiers show the prediction accuracy > 95% in DMR dataset.

5. Moreover like accuracy, the sensitivity and the specificity of the proposed system are also very high in both the datasets.

6. Thus, in both the DBT-TU-JU and DMR datasets, ANN and SVM with polynomial kernel give the highest classification results with two SVs.

However, while doing the performance analysis of the SVs in breast abnormality prediction, it is obvious to evaluate the prediction accuracy of the system with the highly correlated image complexity features so that we can investigate whether it is possible to substitute the SVs with these image complexity features in the proposed system. To do so, we have evaluated the prediction performance of the image complexity features and listed the accuracy values in both datasets by using the individual complexity features and also by using all three complexity features together in Table 5. Like the SVs, the complexity features are also normalized in the same way before feeding them to the classification algorithms. The best classification accuracy of each classifier is represented in *boldface* in Table 5. The listed accuracy value of each classifier against each feature set is the maximum of all the 20 accuracy values obtained in 20 consecutive iterations. As illustrated in Table 5, it has been seen that compare to the SI and FD, the entropy is more discriminative and it provides more prediction accuracy with all classifiers than the SI and FD. And while all complexity features are used in combined way, they provide satisfactory performance with all the classifiers. However, when the performances of these complexity features with all classifiers are compared with the classification performances of the set of 2 SVs (as demonstrated in Table 5), it has been seen that the classification accuracy of the SVs with all classifiers is much better than the classification accuracies of the complexity features. Thus, the SVs are more discriminative in nature than that of the image complexity features.

Table 4 The accuracy, sensitivity and specificity values of each classifier with different number of SVs

Classifiers	DBT-TU-JU						DMR					
	Number of singular values						Number of singular values					
	1	2	5	10	20	30	1	2	5	10	20	30
KNN												
Acc.	89.17	**90.83**	**90.83**	90.00	89.17	88.33	**96.67**	**96.67**	96.56	94.00	93.10	93.81
Sens.	85.71	91.43	88.57	90.00	88.57	88.57	94.00	94.00	90.89	82.22	82.22	80.00
Spec.	94.00	91.00	96.00	90.00	90.00	88.00	98.00	98.00	100.00	99.00	98.00	100.00
SVM_Linear												
Acc.	90.83	**91.67**	90.83	90.00	90.00	89.17	96.00	**98.00**	96.00	95.86	95.20	95.20
Sens.	85.71	91.40	88.57	87.14	87.14	87.14	90.00	96.00	94.00	93.33	91.11	91.11
Spec.	98.00	92.00	94.00	94.00	94.00	92.00	99.00	99.00	97.00	97.00	97.00	97.00
SVM_Poly												
Acc.	91.67	<u>**92.50**</u>	90.83	89.17	89.17	85.00	96.00	<u>**98.00**</u>	96.00	93.90	93.76	93.05
Sens.	88.57	90.00	88.57	87.14	87.14	81.43	94.00	98.00	93.33	91.11	88.89	88.89
Spec.	96.00	96.00	94.00	92.00	92.00	90.00	97.00	98.00	98.00	95.00	96.00	95.00
SVM_RBF												
Acc.	**90.00**	90.00	90.00	89.17	89.17	88.33	**97.33**	96.00	94.67	94.52	94.38	91.76
Sens.	88.57	85.71	87.14	87.14	90.00	88.57	92.00	90.00	86.67	86.67	88.89	84.44
Spec.	92.00	96.00	94.00	92.00	88.57	88.00	100.00	99.00	98.00	98.00	97.00	95.00
DT												
Acc.	90.83	**90.83**	89.17	87.50	88.33	87.50	**95.33**	**95.33**	**95.33**	94.52	94.52	93.14
Sens.	90.00	92.00	90.00	85.71	88.57	88.57	92.00	92.00	92.00	91.11	91.11	91.11
Spec.	90.00	90.00	88.00	90.00	88.00	86.00	97.00	97.00	97.00	96.00	96.00	94.00
ANN												
Acc.	92.50	<u>**92.50**</u>	91.70	91.70	91.70	89.20	97.30	<u>**98.00**</u>	98.00	96.70	96.70	95.20
Sens.	90.00	91.40	91.40	91.40	88.60	85.70	94.00	98.00	96.00	92.00	92.00	84.40
Spec.	96.00	94.00	92.00	92.00	96.00	94.00	99.00	98.00	99.00	99.00	99.00	100.00
RF												
Acc.	88.33	**90.83**	90.83	89.71	88.33	87.50	**96.67**	**96.67**	96.67	96.48	95.90	95.14
Sens.	85.71	91.43	90.00	90.00	91.43	88.57	94.00	94.00	93.33	93.33	88.89	91.11
Spec.	94.00	91.00	92.00	88.00	84.00	86.00	98.00	98.00	98.00	98.00	99.00	97.00
LDA												
Acc.	89.17	90.00	**90.83**	87.50	87.50	85.00	**90.00**	**90.00**	90.00	89.10	88.38	88.29
Sens.	82.85	82.85	85.71	80.00	80.00	75.71	70.00	70.00	68.00	64.44	62.22	64.44
Spec.	98.00	100.00	98.00	98.00	98.00	98.00	100.00	100.00	100.00	100.00	100.00	99.00
AB												
Acc.	**91.67**	90.00	89.17	88.33	88.33	86.67	96.00	**96.67**	96.00	95.90	95.86	95.14
Sens.	88.57	88.57	88.00	87.14	88.57	87.14	94.00	94.00	93.33	93.33	91.11	91.99
Spec.	96.00	92.00	90.00	90.00	88.00	86.00	97.00	98.00	98.00	97.00	98.00	97.00

Boldface—best performance of each classifier, boldface with underline—the highest classification accuracy of the proposed system

Assessment of the grading performance of BAG index

This section evaluates the grading performance of the BAG index. The degree of abnormality or severity that a thermogram bears can be predicted by observing the magnitude of the BAG index. By using Eq. (13), the BAG index values for all breast thermograms have been computed and analyzed to obtain the BAG index range in various categories of thermograms. The boxplots of BAG index values for the normal (N), MA and SA thermograms have been plotted in Fig. 8. As depicted in Fig. 8, the dominant range of BAG index for N, MA and SA thermogram is 0–2, 16–22 and 36–55 respectively and it clearly demonstrates the discriminative capacity of BAG index in the degree of severity prediction. Along with the boxplots, the mean BAG index value of each

Table 5 The prediction accuracy of each classifier with the image complexity measures

Features	Classifiers								
	KNN	SVM_L	SVM_P	SVM_R	DT	ANN	RF	LDA	AB
DBT-TU-JU									
SV_2	**90.83**	**91.67**	**92.50**	**90.00**	**90.83**	**92.50**	**90.83**	**90.83**	**91.67**
Ent	85.83	89.17	89.17	88.33	83.33	88.30	77.50	88.33	85.00
SI	77.50	78.33	78.33	78.33	75.83	74.20	74.17	80.83	75.83
FD	71.67	69.17	68.33	69.10	67.50	76.50	64.17	70.00	70.00
Ent + SI + FD	75.83	74.17	75.00	74.17	89.17	73.50	86.67	75.00	88.33
DMR									
SV_2	**96.67**	**98.00**	**98.00**	**97.33**	**95.33**	**98.00**	**96.67**	**90.00**	**96.67**
Ent	86.86	87.81	89.62	87.41	87.57	89.00	84.05	86.90	89.86
SI	69.52	74.43	72.95	73.00	68.29	73.80	66.33	73.71	73.86
FD	86.48	61.67	65.44	64.10	84.90	74.50	82.48	85.00	68.14
Ent + SI + FD	89.57	74.17	75.00	72.50	91.10	65.50	89.14	88.95	91.05
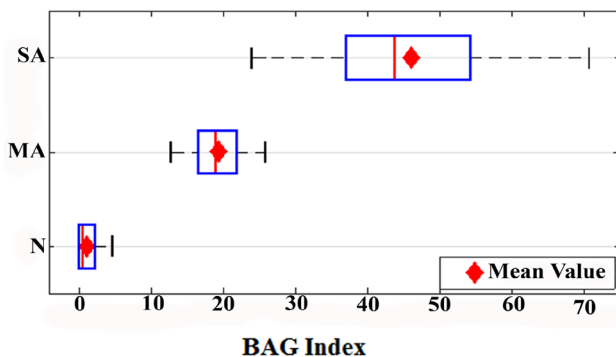
Fig. 8 Boxplots of BAG index values for normal (N), mild abnormal (MA) and severely abnormal (SA) breast thermograms of DBT-U-JU database

group is also plotted in Fig. 8. From the mean value also, we can conclude that each class of thermogram maintain a different range of BAG index values.

Discussion

The traditional approach of breast abnormality prediction is the bilateral asymmetry analysis whose accuracy relies on the proper separation of left and right breasts from a thermogram. The proposed method on the other hand uses the characteristic thermal patterns within the breast regions, whose presence signifies the presence of breast abnormality. The experimental results show the efficiency of our proposed method both in public and in-house datasets. However, to further evaluate the potentiality of the proposed method, this section provides a comparison of our proposed method with other existing works. The comparison of the proposed method is performed in two ways: first, the proposed method is compared with the conventional method of

breast abnormality detection known as bilateral asymmetry analysis as reported in [40] and then, a comparison of the proposed method with other existing works has been made. Along with the comparison, this section also incorporates the limitations of our proposed method.

Comparison of the proposed method with conventional method of breast abnormality detection

In this section, we compare the proposed method with our previously reported method of breast abnormality detection [40]. The breast abnormality detection method reported in [40] was based on the bilateral asymmetry analysis. For doing the bilateral asymmetry analysis, the extracted breast regions were further segmented into left and right breasts. Then, from each breast, a set of 7 first order statistical features: mean, entropy, skewness, kurtosis, variance, standard deviation, maximum intensity value and 17 second-order texture features: contrast, correlation, dissimilarity, energy, entropy, sum entropy, difference entropy, homogeneity, variance, sum of variance, difference variance, autocorrelation, sum average, information measure of correlation1 and information measure of correlation2 were computed. The statistical significance of these extracted features was carried out by using non-parametric Mann–Whitney–Wilcoxon (MWW) test [41] with significance level of 0.1%. Out of these 24 features (7 statistical and 17 texture), 13 and 16 features were found to be statistically significant ($p < 0.001$) for the breast thermograms of DBT-TU-JU and DMR databases respectively. Then considering these statistically significant features, three different feature sets namely: first order statistical features (FStat), second order texture features (STex) and statistically significant (SSigST) features had been formed. Then, the performance of these three feature

sets were evaluated by using different classifiers. Based on the classification performances, it was found that with all the classifiers the statistically significant features gave better performance than both the FStat and STex feature sets in both the datasets. So, here we compare the efficiency of these three feature sets with the proposed breast abnormality detection system to evaluate the superiority of the proposed method over these three sets of features. Since it is aforementioned that the highest classification accuracy is obtained with only two SVs, here for comparison we consider the results obtained using only two SVs.

For comparison, the classification accuracies obtained in 20 iterations with two SVs are statistically evaluated with the classification accuracies of each of FStat, STex, and SSigST feature set in all 20 iterations. Along with the FStat, STex and SSigST feature sets, the classification accuracies of Ent, SI, FD and ESF (Combination of Ent, SI and FD) feature sets in each iteration is also statistically tested. The p-values obtained using the non-parametric MWW test with significance level of 1% and alternative hypothesis, Ha: The classification accuracy of SVs is higher than the classification accuracy of Ent/ SI/ FD/ ESF/ FStat/ STex/ SSigST have been listed in Table 6. With the p-values < 0.01 for each pair of feature set (SVs vs. Ent/SI/FD/ESF/FStat/STex/SSigST) and for each classifier, the MWW test accepts the alternative hypothesis Ha and thus, it can be concluded that for each classifier, the classification accuracy of SVs obtained in each iteration is significantly ($p < 0.01$) greater than the classification accuracies obtained with all other feature sets. Thus, the SVs are found to provide the highest classification accuracy among all the feature sets.

Besides the statistical significance test, the receiver operating characteristics (ROC) curves for all these feature sets with the classifiers that give the highest classification performance have also been plotted in Fig. 9 and the area under the ROC curve (AUC) is considered for the efficiency measure of the feature sets. Along with the ROC curves for SVs, first order, second order and statistically significant features; the ROC curves for the image complexity features are also plotted here for comparison. As reflected in Fig. 9, it has been seen that for both the ANN and SVM_Polynomial classifiers, the SVs return the higher AUC values than the remaining feature sets in both the DBT-TU-JU and DMR datasets, which indicates the superiority of the SVs over the other feature sets. However, even though the classification accuracies of the ANN and SVM_Polynomial are same, the AUC values of SVM_Polynomial are higher than the AUC values of the ANN in both the datasets. In a disease diagnosis system, the higher AUC value (AUC > 0.9) indicates the better prediction accuracy of the system [42, 43]. Hence, by considering both the accuracy and AUC values, it can be concluded that the prediction performance of SVM with polynomial kernel is better than the ANN in both the datasets.

Table 6 Statistical significance of the classification accuracies of SVs against the accuracies of other feature sets

| Classifiers	Features													
	DBT-TU-JU database							DMR database						
	Ent	SI	FD	ESF	FStat	STex	SSigST	Ent	SI	FD	ESF	FS	STex	SSigST
KNN	1.7e−04	2.6e−08	2.4e−08	2.5e−08	2.3e−08	2.6e−08	2.7e−08	3.3e−08	3.3e−08	2.9e−08	3.2e−08	3.3e−08	3.3e−08	3.3e−08
SVM_L	2.5e−08	5.3e−08	1.5e−08	2.4e−08	2.3e−08	2.4e−08	2.4e−08	3.1e−08	3.2e−08	3.2e−08	2.9e−08	3.2e−08	3.2e−08	3.2e−08
SVM_P	1.5e−05	2.2e−08	3.4e−08	1.9e−08	2.6e−08	2.5e−08	2.5e−08	3.2e−08	3.2e−08	1.7e−08	2.4e−08	3.2e−08	3.2e−08	3.2e−08
SVM_R	2.1e−04	1.4e−08	1.6e−08	2.1e−08	2.1e−08	2.1e−08	2.1e−08	3.1e−08	3.1e−08	3.3e−08	2.8e−08	3.3e−08	3.3e−08	3.3e−08
DT	8.6e−04	3.1e−08	3.9e−08	4.2e−07	3.1e−08	3.1e−08	3.1e−08	3.3e−08	3.3e−08	3.2e−08	3.4e−08	3.4e−08	3.4e−08	3.4e−08
ANN	1.8e−08	2.1e−08	2.5e−08	2.3e−08	2.6e−08	2.6e−08	2.5e−08	2.3e−08	2.6e−08	2.6e−08	2.6e−08	2.7e−08	2.7e−08	2.6e−08
RF	2.9e−08	3.1e−08	2.6e−08	7.4e−05	3.1e−08	3.1e−08	3.1e−08	3.4e−08	3.4e−08	3.1e−08	3.4e−08	3.3e−08	3.4e−08	3.4e−08
LDA	4.3e−04	2.1e−08	2.1e−08	2.9e−08	2.5e−08	2.7e−08	2.5e−08	3.2e−08	3.3e−08	3.7e−08	5.6e−08	3.3e−08	3.3e−08	3.3e−08
AB	3.1e−08	2.3e−08	2.7e−08	2.9e−08	2.8e−08	2.8e−08	2.8e−08	3.3e−08	3.3e−08	3.2e−08	3.3e−08	3.3e−08	3.4e−08	3.3e−08
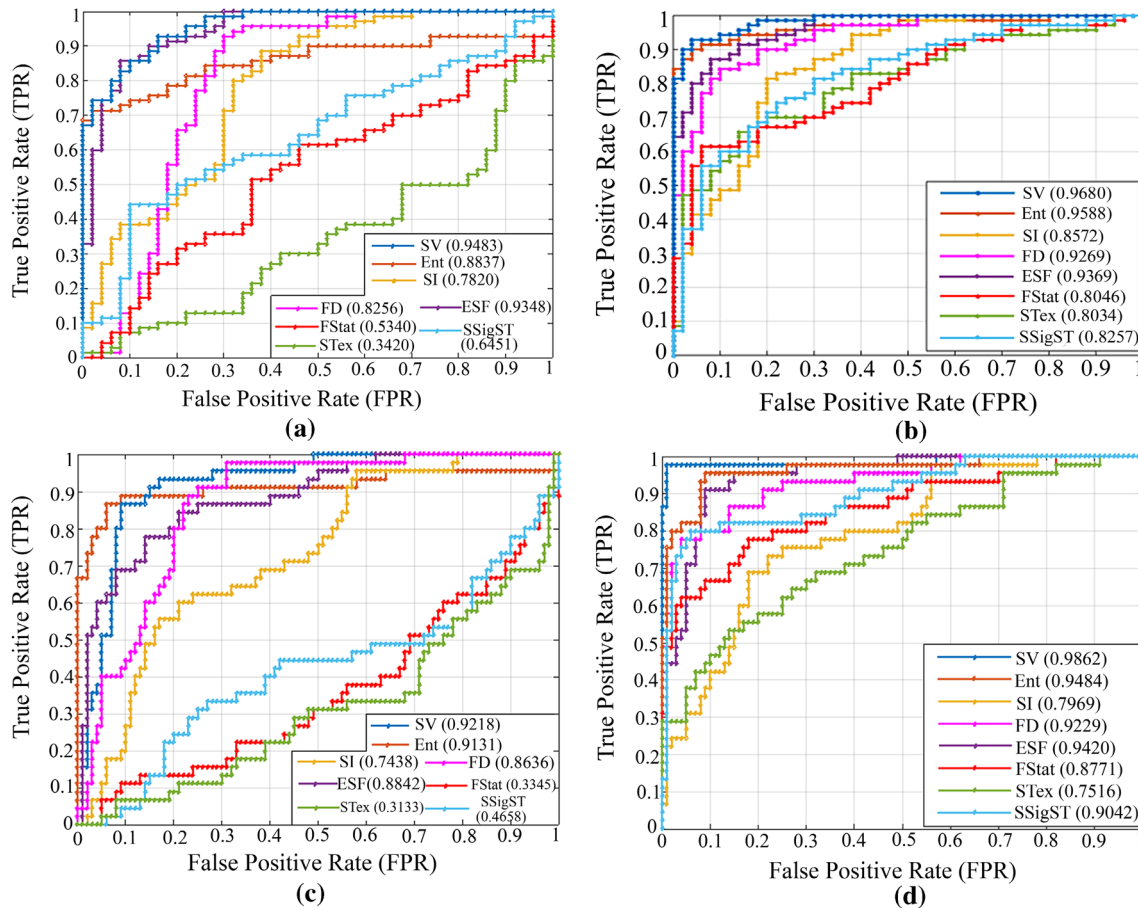
Fig. 9 **a, b** The ROC curves for ANN and SVM_P obtained using the DBT-TU-JU dataset and **c, d** the ROC curves for ANN and SVM_P obtained using the DMR dataset

Comparison with other reported methods

In this section, the comparison of the proposed method is made with the previously reported IBT based breast abnormality detection methods. As described in the literature section of this article, several bilateral asymmetry analysis based breast abnormality detection methods with different degree of efficiency had already been reported. However, it is worth mentioning that due to the unavailability of the breast thermogram databases used in existing research works, we are not able to provide the comparative study of our proposed method in those respective databases. But, there are some bilateral asymmetry analysis based breast abnormality detection methods whose performances were evaluated by using the publically available DMR database. Hence, for comparison purpose, we have also evaluated our proposed system in the publically available DMR database and report the same in Table 7.

As summarized in Table 7, the proposed method outperforms the other existing techniques. Unlike the bilateral asymmetry analysis based methods, the key advantages of

the proposed method is that its accuracy does not rely on the accurate separation of left and right breasts of a thermogram, instead it uses the SVs to characterize the candidate thermal patches whose presence within breast region indicates the breast anomaly. Moreover in contrary to the bilateral asymmetry analysis, the proposed method can detect the presence of breast abnormality even when the abnormality is present in both breasts of a thermogram and thus, improves the accuracy by increasing the number of true positives. Above all, instead of using a large number of features, the proposed method can detect the presence of breast abnormality by using only two SVs and based on these two SVs, it can also predict the degree of abnormality of any thermogram.

Limitation of our proposed method and future work

Although with regard to the obtained performance measures, the proposed method outperforms the existing methods, the ranges of the BAG index for predicting the degree of abnormality is dataset dependent for which the BAG index range for normal, mild abnormal and severely abnormal

Table 7 Prediction performances of the proposed and other reported IBT-based breast abnormality detection methods

Authors	Method used	Dataset	Classifier used	Performance measures		
				Acc.	Sens.	Spec.
Araujo et al. [21]	Temperature based interval symbolic features	DMR: 50 (14M, 19B, 17Cy)	Distance based classifier	–	85.70%	86.50%
Garduno-Ramon et al. [22]	Temperature features	DMR: 454 (414N, 40Ab)	–	–	86.84%	89.43%
Gaber et al. [23]	Statistical, Gabor, texture features	DMR: 63 (29 N, 34M)	SVM	Average Acc. of four scenarios: 88.41%		
Sathish et al. [25]	Statistical and texture features	DMR: 80 (40N, 40Ab)	SVM	90.00%	87.50%	92.50%
Borchartt et al. [26]	Range temperature, mean temperature, standard deviation, quantization of higher tone	DMR: 28	SVM	85.71%	95.83%	25.00%
Our proposed method	1st and 2nd rank singular values	DMR (45Ab, 100N)	SVM (polynomial)	98.00%	98.00%	98.00%

Acc. accuracy, *Sens.* sensitivity, *Spec.* specificity, *N* normal, *B* benign, *C* cancer, *M* malignant, *Cy* cyst, *L* left, *R* right, *Ab* abnormal

may slightly vary depending on the datasets. Moreover, the accuracy of the proposed method depends on the delineation of breast regions from the thermograms so that the thermal patches within the breast region gets characterized but not the thermal patches of the non-breast regions. But, in our proposed method, the delineation of breast region is done by using a semi-automatic segmentation method that needs human-interaction for selection of lower parabolic boundaries. Hence, a potential area for future improvement of our proposed method is to delineate the breast regions automatically from breast thermogram images without the human intervention.

Conclusion

This paper describes a new approach for computer-aided diagnosis of breast abnormality in asymptomatic patients. The task is to automatically distinguish the abnormal thermograms from the normal one to identify the cases that require urgent medical attention. Due to the fact that the thermograms having physiological dysfunction bear more high-temperature regions than that of the healthy breast thermograms, the proposed approach is based on the analysis of the thermal patterns of the breast thermograms. For this purpose, the SVD is used for characterization of these thermal patterns and to distinguish the abnormal thermograms. The experimental results show that the proposed system performs better than other recently reported breast abnormality detection methods. In this paper, our contribution is twofold. First, we have used a single channel of the color channels instead of the three channeled RGB image to represent the candidate thermal patches and then, we have proposed to use the SVs to characterize the thermal patterns and it has brought a remarkable improvement in breast abnormality detection accuracies. We also demonstrate that although the image complexity features are highly correlated with the SVs, the SVs are more efficient in distinguishing the abnormal thermograms from the normal ones. Secondly, based on the extracted SVs, a BAG index has been designed to investigate the degree of abnormalities these abnormal thermograms bear.

Compliance with ethical standards

Conflict of interest All authors declare that they don't have any conflict of interest.

Ethical approval This work is done by maintaining the ethical standards of AGMC with IRB approval number F.4 (5-2)/AGMC/Academic/Project/Research/2007/Sub-I/ 8199-8201.

References

1. Statistics of Breast Cancer in India. http://www.breastcancerindia.net/statistics/trends.html. Accessed 30 Aug 2017

2. Ng EYK (2009) A review of thermography as promising non-invasive detection modality for breast tumor. Int J Therm Sci 48(5):849–859

3. Gogoi UR, Majumdar G, Bhowmik MK, Ghosh AK, Bhattacharjee D, Breast abnormality detection through statistical feature analysis using infrared thermograms. In: Proceeding of IEEE international symposium on advanced computing and communication (ISACC), 2015, pp 258–265

4. Lahiri BB, Bagavathiappan S, Jayakumar T, Philip J (2012) Medical applications of infrared thermography: a review. Infrared Phys Technol 55(4):221–235

5. Schaefer G, Závišek M, Nakashima T (2009) Thermography based breast cancer analysis using statistical features and fuzzy classification. Pattern Recogn 42(6):1133–1137

6. Faria FAC, Cano SP, Gomez-Carmona PM, Sillero M, Neiva CM (2012) Infrared thermography to quantify the risk of breast cancer. Bioimages 20(0):1–7

7. Uematsu S (1985) Symmetry of skin temperature comparing one side of the body to the other. Thermology 1(1):4–7

8. Kwok J, Krzyspiak J (2007) Thermal imaging and analysis for breast tumor detection. Computer-aided engineering: applications to biomedical processes, BEE 453

9. Rastgar-Jazi M, Mohammadi F (2017) Parameters sensitivity assessment and heat source localization using infrared imaging techniques. Biomed Eng Online 16(1):113

10. Qi H, Snyder WE, Head JF, Elliott RL (2000) Detecting breast cancer from infrared images by asymmetry analysis. In: Proceedings of the 22nd IEEE annual international conference of the IEEE Engineering in Medicine and Biology Society, vol 2, pp 1227–1228

11. Borchartt TB, Conci A, Lima RCF, Resmini R, Sanchez A (2013) Breast thermography from an image processing viewpoint: a survey. Signal Process 93(10):2785–2803

12. Gogoi UR, Bhowmik MK, Bhattacharjee D, Ghosh AK, Majumdar G (2016) A study and analysis of hybrid intelligent techniques for breast cancer detection using breast thermograms. In: Bhattacharyya S, Dutta P, Chakraborty S (eds) Hybrid soft computing approaches. Springer, New Delhi, pp 329–359

13. Bhowmik MK, Gogoi UR, Majumdar G, Datta D, Ghosh AK, Bhattacharjee D (2018) Designing of ground truth annotated DBT-TU-JU breast thermogram database towards early abnormality prediction. IEEE J Biomed Health Inform (JBHI) 22(4):1238–1249

14. Silva LF, Saade DCM, Sequeiros-Olivera Silva GO, Paiva AC, Bravo RS, Conci A (2014) A new database for breast research with infrared image. J Med Imaging Health Inform 4(1):92–100 9

15. Ng EYK, Kee EC (2007) Integrative computer-aided diagnostic with breast thermogram. J Mech Med Biol 7(01):1–10

16. Mookiah MRK, Acharya UR, Ng EYK (2012) Data mining technique for breast cancer detection in thermograms using hybrid feature extraction strategy. Quant Infrared Thermogr J 9(2):151–165

17. Acharya UR, Ng EYK, Tan JH, Sree SV (2012) Thermography based breast cancer detection using texture features and support vector machine. J Med Syst 36(3):1503–1510

18. Francis SV, Sasikala M (2013) Automatic detection of abnormal breast thermograms using asymmetry analysis of texture features. J Med Eng Technol 37(1):17–21

19. Francis SV, Sasikala M, Bharathi GB, Jaipurkar SD (2014) Breast cancer detection in rotational thermography images using texture features. Infrared Phys Technol 67:490–496

20. Francis SV, Sasikala M, Saranya S (2014) Detection of breast abnormality from thermograms using curvelet transform based feature extraction. J Med Syst 38(4):23

21. Araújo MC, Lima RC, De Souza RM (2014) Interval symbolic feature extraction for thermography breast cancer detection. Expert Syst Appl 41(15):6728–6737

22. Garduño-Ramón MA, Vega-Mancilla SG, Morales-Henández LA, Osornio-Rios RA (2017) Supportive noninvasive tool for the diagnosis of breast cancer using a thermographic camera as sensor. Sensors 17(3):497

23. Gaber T, Ismail G, Anter A, Soliman M, Ali M, Semary N, Snasel V (2015) Thermogram breast cancer prediction approach based on neutrosophic sets and fuzzy c-means algorithm. In: 37th IEEE annual international conference in Engineering in Medicine and Biology Society (EMBC), pp 4254–4257

24. Zadeh HG, Haddadnia J, Hashemian M, Hassanpour K (2012) Diagnosis of breast cancerusing a combination of genetic algorithm and artificial neural network in medical infrared thermal imaging. Iran J Med Phys 9(4):265–274

25. Sathish D, Kamath S, Prasad K, Kadavigere R, Martis RJ (2017) Asymmetry analysis of breast thermograms using automated segmentation and texture features. Signal Image Video Process 11(4):745–752

26. Borchartt TB, Resmini R, Conci A, Martins A, Silva AC, Diniz EM, Lima RC (2011) Thermal feature analysis to aid on breast disease diagnosis. In: Proceedings of 21st Brazilian Congress of Mechanical Engineering—COBEM2011, pp 24–28

27. Lashkari A, Pak F, Firouzmand M (2016) Full intelligent cancer classification of thermal breast images to assist physician in clinical diagnostic applications. J Med Signals Sens 6(1):12

28. Suganthi SS, Ramakrishnan S (2014) Semi-automatic segmentation of breast thermograms using variational level set method. In: Proceedings of 15th international conference on biomedical engineering, Springer International Publishing, Singapore, pp 231–234

29. Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios. Recommendation ITU-R BT.601-7, 2011

30. Luo H, Lin D, Yu C, Chen L (2013) Application of different HSI color models to detect fire-damaged mortar. Int J Transp Sci Technology 2(4):303–316

31. Gonzalez RC, Woods RF (1992) Digital image processing. Addison Wesley, Reading

32. Matrix Norms. http://fourier.eng.hmc.edu/e161/lectures/algebra/node12.html. Accessed 11 April 2018

33. Petersen KB, Pedersen MS (2008) The matrix cookbook, vol. 7, no. 15. Technical University of Denmark, Kongens Lyngby, pp 510

34. Matrix Norms and Condition Numbers. http://faculty.nps.edu/rgera/MA3042/2009/ch7.4.pdf. Accessed 11 April 2018

35. Veisi H, Jamzad M (2009) A complexity-based approach in image compression using neural networks. Int J Signal Process 5(2):82–92

36. Yu H, Stefan W (2013) Image complexity and spatial information. In: Quality of Multimedia Experience (QoMEX), fifth international workshop on Klagenfurt am Wörthersee, IEEE, pp 12–17

37. Backes AR, Bruno OM (2010) Medical image retrieval based on complexity analysis. Mach Vision Appl 21(3):217–227

38. Falconer K (2004) Fractal geometry: mathematical foundations and applications. Wiley, New York

39. Fisher RA (1934) Statistical methods for research workers, 13th edn. Hafner, New York

40. Gogoi UR, Bhowmik MK, Ghosh AK, Bhattacharjee D, Majumdar G (2017) Discriminative feature selection for breast abnormality detection and accurate classification of thermograms. In: Proceeding of IEEE international conference on innovations in electronics, signal processing and communication (IESC), pp 39–44

41. Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics 1(6):80–83

42. Karimollah HT (2013) Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J Intern Med 4(2):627–635

43. Mandrekar JN (2010) Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 5(9):1315–1316