Discriminative Feature Selection for Breast Abnormality Detection and Accurate Classification of Thermograms

Usha Rani Gogoi[1], Mrinal Kanti Bhowmik[1], Anjan Kumar Ghosh[1], Debotosh Bhattacharjee[2], Gautam Majumdar[3]
[1]Department of Computer Science & Engineering, Tripura University (A Central University), Suryamaninagar-799022, Tripura
[2]Department of Computer Science & Engineering, Jadavpur University, Kolkata-700032, West Bengal
[3]Radiotherapy Department, Regional Cancer Center, Agartala Government Medical College, Agartala-799006, Tripura
ushagogoi.cse@gmail.com, mrinalkantibhowmik@tripurauniv.in, anjn@ieee.org, debotosh@ieee.org, drgmajumdar@yahoo.in

***Abstract*—Infrared breast thermography with the potential of predicting the future risk of developing breast cancer, has been considered as an early breast abnormality detection tool. This paper investigates the importance of selecting the discriminative features for improving the classification accuracy of the infrared thermography based breast abnormality detection systems. Mann-Whitney-Wilcoxon statistical test has been used here to select the best discriminative features from a feature set of 24 features, extracted from each breast thermogram of DBT-TU-JU and DMR databases. Three set of features: FStat, STex and SSigFS generated from these 24 extracted features are then fed into six most widely used classifiers for comparing the efficiency of each feature set in breast abnormality detection. The experimental results show that among all three feature sets, statistically significant feature set (SSigFS) provides more accuracy in discriminating the abnormal breast thermograms from the normal.**

***Index Terms*—Infrared Breast Thermography, Discriminative features, Feature selection, Classification.**

I. INTRODUCTION

Detection and diagnosis of breast diseases before the onset of a malignant tumor is the only way to reduce the incidence rate of breast cancer. Infrared breast thermography with the capability of identifying the presence of breast abnormality before the onset of any malignant tumors has been adopted as a breast health monitoring method. It has the potentiality for signifying the future risk of developing breast cancer [1]. With the advancement of infrared thermography and image processing techniques, several research works have been carried out for infrared thermography based early breast abnormality detection. Among various breast abnormality detection methods, analysis of asymmetric thermal patterns in bilateral breasts by using statistical and texture features is found to be the most common and efficient method for the early breast abnormality detection [2], [3]. In our previous work [3] [4], a survey on statistical and texture feature based asymmetry analysis of the breast thermograms was made. In most of the works, the experimental results showed the effectiveness of different statistical and texture features along with different classifiers in early breast abnormality detection and classification of thermograms. However, sometimes use of redundant features in breast thermogram classification reduces the accuracy of the breast abnormality prediction systems. In this perspective, we emphasize to find the necessity of selecting the discriminative features for improving the prediction accuracy of a classifier system.

In this work, for predicting the presence of any abnormality in breast thermograms, a set of 24 features comprising of 7 statistical and 17 texture features has been extracted from both breasts of a thermogram that quantifies the symmetric and asymmetric thermal patterns of the breasts. Extraction of the features is followed by the selection of the discriminative features by doing Mann-Whitney-Wilcoxon (MWW) test with $p<0.005$. Then, three feature sets: FStat (First order statistical features), STex (Second-order Texture Feature) and SSigFS (Statistically significant first order and second order features) are made and their potentiality in abnormality detection is measured by using six most widely used classifier systems. The paper contributes by investigating the importance of feature selection for classification of breast thermograms which is very crucial to reduce the computational burden of any breast abnormality detection system. Moreover, it explores the most proficient combination of feature set and classifier that can provide better accuracy in breast abnormality detection.

The paper is organized as follows. Section II provides a brief description of the experimental datasets: DBT-TU-JU and DMR. Description of the datasets is followed by the analysis of breast thermograms in Section III, which involves the segmentation of breast thermograms, temperature analysis of breast thermograms, statistical and texture feature based asymmetry analysis of breast thermograms and selection of best discriminative features. Section IV compares the classification accuracy of three different feature sets with six

978-1-5090-5620-0/17/$31.00 ©2017 IEEE

commonly used classifiers to illustrate the importance of feature selection for breast abnormality prediction. Finally, Section V concludes the paper.

II. Database Description

A. DBT-TU-JU Breast Thermogram Database

The DBT-TU-JU is our own created breast thermogram database developed at the Regional Cancer Centre (RCC), Agartala Government Medical College (AGMC), Govt. of India. The database is designed under strict acquisition protocols discussed in [5] using the FLIR T650sc thermal camera with a resolution of 640 x 480 pixels and thermal sensitivity of <20 mK at 30^{o}C. In order to use the breast thermograms in research purpose, the written consent of each subject was also taken. Currently, the database comprises of the breast thermograms of 70 patients, out of which 46 breast thermograms are collected from medically proven unhealthy patients and 24 breast thermograms are taken from medically proven healthy subjects. For the purpose of experiment, the entire dataset has been used.

B. DMR Breast Thermogram Database

The DMR (Database for Mastology Research) Database [6] is a publically available breast thermogram database. The database contains the breast thermograms of total 287 subjects out of which, the thermograms of 47 subjects are labeled as 'Sick' and remaining thermograms are labeled as 'Healthy'. For acquisition of the thermograms, they had used FLIR SC-620 Thermal Camera with a spatial resolution of 640 x 480 pixels. For experimental purpose, 35 abnormal and 45 normal breast thermograms are randomly selected.

III. Analysis of Breast Thermograms

A human body is thermally symmetrical, i.e., the temperature distribution of left part of the body is almost symmetrical to the temperature distribution of the right part [7]. Based on this key idea, the presence of any bilateral asymmetry or asymmetric thermal pattern in a breast thermogram may provide some clues to the breasts' pathological conditions. The formation of a cancerous breast tumor is associated with high metabolic activities that cause an increase in local temperature in comparisons to the normal cells [8]. Being a functional imaging modality, infrared breast thermography does not give any structural information (like tumor size, architectural distortion, and micro-calcifications, etc.). It is just a mapping of the skin temperature or infrared energy emitted by the human body parts. Hence, one widely used method for breast abnormality detection is to make a comparison between the thermal patterns of the left and right breasts which is also known as the bilateral asymmetry analysis.

A. Breast Region Segmentation from Breast Thermograms

In the analysis of breast thermograms, the proper segmentation of the breast region plays a vital role. The primary objective of segmenting a breast thermogram is to discard all those portions which are not belonging to the breast region. However, the inframammary folds and unclear lower breast boundaries make the automatic segmentation more complicated. The segmentation of breast thermogram can either be manual or automatic. In our datasets, in most of the cases, the lower parabolic boundaries of both breasts are not prominent enough to be detected automatically. But, while doing segmentation of breast thermograms, it is very crucial that the breast areas in lower breast boundaries and inframammary folds are not getting lost [9]. Hence, in this work we are adopting a semi-automatic method for segmenting the breast region, which requires the human interaction for selecting the lower parabolic boundaries of both breasts. The algorithm for segmentation of breast region from a breast thermogram is as follows. To perform the segmentation of breast region, initially the breast images are manually cropped to remove the unnecessary areas including the neck portion, the area underneath breasts and the background. Fig. 1 shows the detailed overview of the segmentation procedure.

1) Convert the images from RGB to YCbCr color space and extract the blue difference Chroma component Cb from the thermogram images.

2) Binarization of the Cb channel to distinguish the body area from the background.

3) Selection of the lower breast boundary along with the upper abdomen portion visible in the thermogram to discard the remaining unnecessary regions.

4) Subtract the selected unnecessary region in Step 3 from the binary image of Step2 to generate the individual breast mask.

5) Convolve the generated mask to the RGB image to get the segmented breast region.

6) Separate the left and right breasts from segmented breast region.

B. Temperature Analysis of Breast Thermograms

In the analysis of breast thermograms, the temperature analysis is one of the essential steps where a small temperature difference between the left and right breast may indicate any breast pathology. Freitas stated that the temperature difference between symmetrical areas should not exceed 0.5 ^{o}C [9]. This information of temperature variation

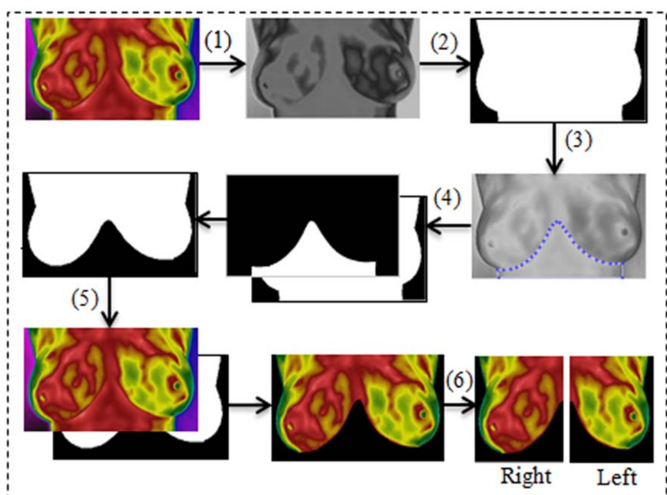

Figure 1. Segmentation of the Breast Region from a Breast Thermogram

can be used to detect the presence of breast abnormality. Temperature analysis is based on the premise that the temperature distribution of one breast of an abnormal breast thermogram significantly varies from the temperature distribution of the other breast. It is because the surface temperature of the skin above the anomalous breast region is considerably higher than an area further away from that anomalous region [10]. However, the amount of heat radiation varies based on the location of the tumor within the breast. Here, from each thermogram of the normal and abnormal group, the mean, mode and maximum temperatures (in ^{o}C) are computed and the differences between the left and right breast values are found out to show the requirement of temperature analysis in breast abnormality prediction. The average of mean, mode and maximum temperature of the normal group and the abnormal group of both the DBT-TU-JU and DMR databases are listed in Table I. From the average temperature values listed in Table I, it has been seen that the average temperature difference of all the three parameters: mean, maximum and mode of the abnormal group are significantly higher than the average temperature difference of the normal group in both databases. In the abnormal group, the average of mean, maximum and modal temperature differences between the left and right breast are found to be greater than or equal to 0.5^{o}C. On the other hand, the averages of these temperature differences are found to be less than 0.3^{o}C in case of the normal group. Thus, the temperature analysis of the thermograms can signify the presence of breast pathology.

C. Statistical & Texture Feature based Analysis of Breast Thermograms

An effective approach to automatically detect breast abnormality is to study the symmetry between the left and the right breast i.e., the bilateral symmetry [11]. To quantitatively represent the symmetric or asymmetric thermal patterns between two breasts, extraction of statistical and texture features are very effective. In other words, these statistical and texture features provide a significant way of representing an image into quantitative measures. Here to detect the breast abnormality, a set of 24 features including both first order statistical features (7) and texture features (17) are extracted from the left and right breast of each breast thermogram. The first order statistical features are intensity histogram based features since they can directly be computed from the probability distribution of each intensity value. The first order features computed in this work are: mean, entropy, skewness, kurtosis, variance, standard deviation and maximum intensity value. To identify the presence of any abnormality in a breast thermogram, the difference of the left and right breast feature values is computed for each thermogram. Similarly, the differences of the 17 texture features are also found out. The texture features computed here are: contrast, correlation, dissimilarity, energy, entropy, sum entropy, difference entropy, homogeneity, variance, sum variance, difference variance, sum average, autocorrelation, Information measure of correlation1 (Inf1), Information measure of correlation2 (Inf2), Inverse difference normalized (IDN), Inverse difference moment normalized (IDMN). After calculating the feature value differences in both normal and abnormal group, the mean and standard deviation of each feature difference is computed and tabularized in Table II. From these feature value differences, we can conclude that compared to a normal breast thermogram, a breast thermogram of the abnormal group possesses significant difference between the two breasts.

TABLE I. THE AVERAGE OF MEAN, MAXIMUM AND MODAL TEMPERATURES (IN ^{0}C) OF THE NORMAL AND ABNORMAL GROUP

DB Used			Average Temperature Values		
			MeanTemp.	***Max. Temp.***	***Modal Temp.***
DBT-TU-JU	Normal	Left	34.74 ± 0.74	32.92 ± 0.68	32.90 ± 1.21
		Right	34.45 ± 0.82	32.68 ± 0.89	32.70 ± 1.05
		Diff.	**0.29 ± 0.22**	**0.240 ± 0.22**	**0.20 ± 0.16**
	Abnormal	Left	35.22 ± 1.07	33.17 ± 1.28	33.35 ± 1.31
		Right	35.80 ± 1.00	32.66 ± 1.27	32.81 ± 1.61
		Diff.	**0.579 ± 0.58**	**0.509 ± 0.50**	**0.54 ± 0.57**
DMR	Normal	Left	28.73 ± 0.85	32.42 ± 1.51	23.94 ± 1.17
		Right	28.55 ± 0.90	32.64 ± 1.36	24.27 ± 1.14
		Diff	**0.178 ± 0.33**	**0.216 ± 0.71**	**0.332 ± 0.45**
	Abnormal	Left	31.27 ± 1.07	34.57 ± 1.02	25.53 ± 0.91
		Right	30.64 ± 0.93	33.90 ± 1.10	26.05 ± 1.56
		Diff	**0.628 ± 0.14**	**0.675 ± 0.84**	**0.524 ± 1.24**

D. Best Discriminative Feature Selection for Breast Abnormality Detection

In the designing of a breast abnormality prediction system, the selection of features that are more discriminative in nature i.e., that can significantly differentiate an abnormal thermogram from a normal thermogram is very important. By discarding the irrelevant and redundant features, whose contribution to the accuracy of the abnormality detection system is very negligible, the feature selection can reduce the complexity of the breast abnormality detection model. Hence here in this work, the extraction of feature values is followed by the statistical test of the extracted feature values that allows us to make an inference from the features. It also allows evaluating the credibility and discriminability of features for breast abnormality detection and hence, plays a very vital role to make appropriate decisions. The statistical test of significance finds out only those features that reach significant difference ($p<0.005$) and these features are considered as the most discriminative features for breast abnormality detection. The significance level of each feature value is measured against the null hypothesis and tabulated in Table II. As illustrated in Table III, in comparison to the normal group, most of the abnormal group feature differences are found to be significantly greater ($p<0.005$) in both the DBT-TU-JU and DMR databases. The p-value of features which reach the significant difference of $p<.005$ are marked as *significant* in the second column of the *Significance level* in Table II. The non-parametric Mann-Whitney-Wilcoxon (MWW) test with significance levels $p<0.005$ has been used to perform the statistical test. The null hypothesis is H_o: The median of the abnormal group is less than the median of the normal group. But, the statistical test in most of the features rejected the null hypothesis by accepting the alternative hypothesis H_a: The median of the abnormal group is greater than the median of the normal group. Based on this statistical test, 13 features computed from the thermograms of the DBT-TU-JU database show the ability to differentiate the abnormal and normal breast thermograms. These 13 best discriminative features include mean, entropy, kurtosis, variance, standard deviation, correlation, higher order entropy, sum entropy, higher order variance, sum of variance, autocorrelation, sum average, Information measure of correlation. Similarly, among the features computed from the thermograms of DMR database, 16 features including mean, entropy, skewness, kurtosis, variance, standard deviation, correlation, dissimilarity, energy, higher order entropy, sum entropy, higher order variance, sum of variance, autocorrelation, sum average and information measure of correlation are found to be more significant in breast abnormality prediction than the remaining features.

IV. Classification of Breast Thermograms

In the designing of a breast abnormality detection system, the classification of thermograms into normal and abnormal group is most crucial. However, choosing the combination of the most efficient feature set and the classifier is one of the most challenging tasks in machine learning. In this section, we investigate the performance of four very commonly used classifiers against three different feature sets for the classification of breast thermograms. Two datasets of 70 breast thermograms (46 Abnormal and 24 Normal) and 80 breast thermograms (35 Abnormal and 45 Normal) from DBT-TU-JU and DMR databases respectively have been used here to make the investigation. The three feature sets: features based on first order statistical features (FStat), second order texture features (STex) and statistically significant first order and second order features (SSigFS) are used to evaluate the performance of four commonly used classifiers: Support Vector Machine (SVM), K-nearest neighborhood (KNN), Artificial Neural Network (ANN) and Decision tree for breast abnormality detection. In case of SVM, instead of using a single kernel to perform the classification, three different kernels are used for evaluation. Thus, a total of six classifiers has been used here to evaluate the performance of feature sets.

The support vector machine (SVM) is the most widely used supervised learning method for classification. It is a powerful methodology to deal with the problem of the unbalanced dataset and nonlinear classification. It can minimize the empirical classification error and maximize the geometric margin of a classifier i.e., the SVM is based on the principle of structural risk minimization [12]. In literature, it has been reported that due to the property of focusing the

TABLE II Average values of statistical and texture Features Obtained from the Normal and Abnormal Group and Their Significance Level for Breast Abnormality Prediction

		DBT-TU-JU Database				DMR Database			
Extracted Features		**Difference (Mean ± SD)**		**Significance level**		**Difference (Mean ± SD)**		**Significance level**	
		Normal (24)	**Abnormal (46)**	**P value (<.005)**		**Normal (45)**	**Abnormal (35)**	**P value (<.005)**	
Statistical Features	Mean	4.72 ± 2.86	12.02 ± 9.27	0.000050463	**Significant**	5.71 ± 4.66	15.82 ± 13.99	0.00002493	**Significant**
	Entropy	0.05 ± 0.03	0.19 ± 0.13	0.000000009	**Significant**	0.04 ± 0.04	0.23 ± 0.19	0.00000012	**Significant**
	Skewness	0.18 ± 0.11	0.33 ± 0.27	0.025400000	--	0.15 ± 0.11	0.57 ± 0.48	0.00000002	**Significant**
	Kurtosis	0.21 ± 0.23	0.49 ± 0.38	0.000151400	**Significant**	0.09 ± 0.06	1.26 ± 1.13	0.00000000	**Significant**
	variance	0.01 ± 0.01	0.02 ± 0.01	0.000001006	**Significant**	0.00 ± 0.00	0.01 ± 0.01	0.00000000	**Significant**
	Std. Deviation	0.00 ± 0.00	0.01 ± 0.01	0.000000934	**Significant**	0.01 ± 0.01	0.03 ± 0.02	0.00000000	**Significant**
	Max	6.67 ± 5.89	16.09 ± 13.90	0.011502533	--	13.14 ± 13.92	19.34 ± 17.54	0.1104	--
Texture Features	Contrast	0.02 ± 0.02	0.02 ± 0.02	0.235956578	--	0.02 ± 0.02	0.03 ± 0.03	0.0127	--
	Correlation	0.01 ± 0.01	0.01 ± 0.01	0.001489075	**Significant**	0.00 ± 0.00	0.01 ± 0.01	0.0014	**Significant**
	Dissimilarity	0.01 ± 0.01	0.01 ± 0.01	0.724466574	--	0.01 ± 0.01	0.02 ± 0.02	0.0035	**Significant**
	Energy	0.01 ± 0.01	0.03 ± 0.02	0.039966618	--	0.02 ± 0.02	0.03 ± 0.03	0.0011	**Significant**
	Entropy	0.06 ± 0.04	0.11 ± 0.06	0.002264234	**Significant**	0.07 ± 0.05	0.18 ± 0.13	0.00003682	**Significant**
	Sum Entropy	0.05 ± 0.04	0.11 ± 0.07	0.000241115	**Significant**	0.06 ± 0.05	0.16 ± 0.13	0.00000260	**Significant**
	Diff. Entropy	0.02 ± 0.02	0.03 ± 0.02	0.646061063	--	0.02 ± 0.02	0.03 ± 0.03	0.0753	--
	Homogeneity	0.01 ± 0.01	0.01 ± 0.01	0.556160364	--	0.00 ± 0.00	0.01 ± 0.01	0.0071	--
	Variance	2.14 ± 1.82	6.33 ± 4.83	0.000006212	**Significant**	1.02 ± 1.14	3.14 ± 2.52	0.00003236	**Significant**
	Sum of variance	0.81 ± 0.70	2.30 ± 1.73	0.000017285	**Significant**	2.64 ± 3.01	8.40 ± 6.58	0.00000439	**Significant**
	Diff. variance	0.02 ± 0.02	0.02 ± 0.02	0.235956578	--	0.02 ± 0.02	0.04 ± 0.03	0.0627	--
	Autocorrelation	0.81 ± 0.70	2.31 ± 1.74	0.000020782	**Significant**	1.15 ± 1.02	3.16 ± 2.52	0.00002967	**Significant**
	Sum average	0.23 ± 0.20	0.55 ± 0.41	0.000457784	**Significant**	0.31 ± 0.29	0.80 ± 0.65	0.00000638	**Significant**
	Inf1	0.02 ± 0.01	0.02 ± 0.02	0.551304737	--	0.01 ± 0.01	0.02 ± 0.02	0.2482	--
	Inf2	0.01 ± 0.00	0.01 ± 0.01	0.000074923	**Significant**	0.00 ± 0.00	0.01 ± 0.01	0.00000887	**Significant**
	Idn	0.00 ± 0.00	0.00 ± 0.00	0.580300066	--	0.00 ± 0.00	0.00 ± 0.00	0.0054	--
	Idmn	0.00 ± 0.00	0.00 ± 0.00	0.498683107	--	0.00 ± 0.00	0.00 ± 0.00	0.0652	--

support vectors (training samples that are most difficult to classify), SVM outperforms the competing methods in many applications [12]. Here, in this work we explore the performance of the SVM classifier with three different kernels: i) Gaussian Radial Basis Function kernels (SVM_RBF), ii) Linear Kernel (SVM_Linear) and iii) Polynomial kernel (SVM_Polynomial). Except for the linear kernel, the parameters of both polynomial and the radial 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.

K-nearest neighborhood (KNN) is a robust but lazy learning algorithm since it does not abstract any information from the training data, i.e., it does not have any explicit training phase for which its training phase is very fast. KNN is instance-based classifier where the classification of an unknown instance is done by relating the unknown instance to some known instances based on distance, i.e., an object is assigned with a class to which majority of its neighbors belongs. Thus for a KNN classifier, it is important to consider only those neighbors whose correct classification is already known to make a correct classification. The higher value of K (Neighbors) makes the system less locally sensitive and improves the accuracy. For, a binary classifier, the value of K should be an odd number [13]. Here, the values of neighbors K is chosen to be 9 for classification, which gives best results compared to the other values of K.

Like the other classification model, the key objective of the Decision tree classifier is to identify a model that best fits the association between the attribute set and class label set of the training data [14]. Decision tree classifier represents their classification knowledge in tree form, especially in binary. They have the ability to select the most discriminatory features. In a decision tree, the root node and the internal nodes contain the attribute test conditions to separate the data that have different characteristics and the terminal node gives the decision about class labels. Thus, decision trees are very flexible and easy to understand since the output of a decision tree can be easily interpreted as rules [14].

Artificial neural network is a computer model inspired by the human neural architecture. The Artificial Neural Network is widely used in pattern recognition problems. It is considered as a robust classifier as they are capable of modeling highly non-linear systems in which the variables' relationships are unknown or complex [15].

The diagnostic performance of a classifier is generally evaluated by factors like accuracy, sensitivity and specificity which are generally computed from the parameters of confusion matrix like True Positive (TP), False Negative (FN), True Negative (TN) and False Positive (FP). 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 samples that are correctly classified. Thus, if both sensitivity and specificity are high (low), the accuracy will also be high (low). But, if any one of the sensitivity or specificity is high, then the accuracy will be biased to any one of these measures. These performance measures of each of the classifiers are measured against the three feature set extracted from the breast thermograms to investigate their prediction performance against each feature set for breast abnormality detection. The accuracy, sensitivity, specificity of each of these classifiers against each feature set is illustrated in Table III.

The following observations can be made from Table III for both the databases:

1) The classification accuracy with the statistically significant features is considerably higher in comparison to both first-order statistical feature set and second order texture feature set for all classifiers.

TABLE III ACCURACY, SENSITIVITY & SPECIFICITY OF EACH OF THE CLASSIFIERS AGAINST EACH FEATURE SET

		DBT-TU-JU Database			**DMR Database**		
Classifiers	**Features**	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
SVM_RBF	FStat	81.43	76.09	91.67	82.50	90.00	75.00
	STex	77.14	78.26	75.00	67.50	65.00	70.00
	SSigFS	**84.29**	82.60	**87.50**	85.00	90.00	80.00
SVM_Linear	FStat	71.43	63.04	87.50	87.50	90.00	85.00
	STex	68.57	60.87	83.33	62.50	65.00	60.00
	SSigFS	75.71	69.57	**87.50**	**87.50**	**95.00**	80.00
SVM_Polynomial	FStat	72.86	76.09	66.67	72.50	60.00	85.00
	STex	58.57	65.22	45.83	57.50	65.00	50.00
	SSigFS	74.29	76.09	70.83	72.50	75.00	70.00
KNN	FStat	71.43	73.91	66.67	70.67	77.14	65.00
	STex	61.43	69.56	45.83	67.44	61.11	72.00
	SSigFS	74.29	73.91	75.00	72.00	85.71	60.00
Decision Tree	FStat	75.71	84.78	58.33	84.00	**100.00**	73.33
	STex	67.14	73.91	54.17	72.00	60.00	80.00
	SSigFS	75.71	84.78	58.33	84.00	**100.00**	73.33
ANN	FStat	80.00	80.43	79.17	85.00	80.00	88.89
	STex	67.14	78.26	45.83	78.00	65.71	88.89
	SSigFS	**84.29**	**89.13**	75.00	**87.50**	80.00	**93.33**

2) In DBT-TU-JU database, in comparison to other classifiers, SVM_RBF and ANN both produce a classification accuracy rate of 84.29% with the statistically significant feature set. However, the sensitivity of ANN is better than the sensitivity of SVM_RBF indicating that ANN is more accurate in correctly identifying the positive cases among all positive cases having positive screening test or severe breast problems.

3) In DMR database, among all the classifiers, SVM_Linear and ANN show the accuracy rate of 87.50 with the statistically significant features. Although the accuracies of these two classifiers are same, but SVM_Linear with sensitivity of 95% is found to be more accurate in breast abnormality prediction in positive cases than the ANN. And, with 93.33% of specificity, ANN is more accurate in predicting the absence of breast abnormality in disease-free objects.

4) In both the DBT-TU-JU and DMR databases, the decision tree shows the same value of accuracy, sensitivity and specificity for both first order features and the statistically significant features which signify that for decision tree, instead of using a larger feature space, the first order feature space is sufficient to obtain the same accuracy result. Moreover, it provides the highest sensitivity of 100% in DMR database indicating its efficiency in predicting the abnormal breast thermograms.

V. Conclusion

This work emphasizes the importance of selecting best discriminative features for improving the accuracy of the classification system in breast abnormality detection. The statistically significant feature set shows better performance in distinguishing the abnormal breast thermograms from the normal in both the DBT-TU-JU and DMR databases. Among all six classifiers, both SVM_RBF and ANN provide the highest classification accuracy of 84.29% in DBT-TU-JU database. On the other hand, in DMR database, along with the SVM_Linear, the ANN also provides highest classification accuracy of 87.50%. Obtaining, better accuracy results with the statistically significant features, it can be concluded that selection of discriminant feature is important to improve the accuracy of an infrared thermography based breast abnormality detection system.

Acknowledgment

The research work is supported by the Grant No. BT/533/NE/TBP/2013, Dated 03/03/2014 from the Department of Biotechnology (DBT), Government of India. The first author is grateful to Department of Science and Technology (DST), Government of India for providing her Junior Research Fellowship (JRF) under DST-INSPIRE fellowship program (No. IF150970).

References

[1] F. A. C. Faria, S. P. Cano, P. M. Gomez-Carmona, M. Sillero, C. M. Neiva, "Infrared thermography to quantify the risk of breast cancer," *Bioimages*, vol. 20, pp. 1-7, 2012.

[2] H. Qi, and F. H. Jonathan, "Asymmetry analysis using automatic segmentation and classification for breast cancer detection in thermograms," in *Proc. 2001 IEEE Annual International Conference on Engineering in Medicine and Biology Society*, pp. 2866-2869.

[3] U. R. Gogoi, G. Majumdar, M. K. Bhowmik, A. K. Ghosh, and D. Bhattacharjee, "Breast abnormality detection through statistical feature analysis using infrared thermograms," in *2015 IEEE Advanced Computing and Communication (ISACC), International Symposium*, pp. 258-265.

[4] U. R. Gogoi, M. K. Bhowmik, D. Bhattacharjee, A. K. Ghosh, and G. Majumdar, "A Study and Analysis of Hybrid Intelligent Techniques for Breast Cancer Detection Using Breast Thermograms," in *Hybrid Soft Computing Approaches*, pp. 329-359, Springer India, 2016.

[5] M. K. Bhowmik, U. R. Gogoi, K. Das, A. K. Ghosh, D. Bhattacharjee and G. Majumdar, "Standardization of infrared breast thermogram acquisition protocols and abnormality analysis of breast thermograms," in *SPIE Commercial + Scientific Sensing and Imaging*, pp. 986115-(1-18), 2016.

[6] L. F. Silva, D. C. M. Saade, G. O. S. O. Silva, A. C. Paiva, R. S. Bravo, A. Conci, "A new database for breast research with infrared image," *Journal of Medical Imaging and Health Informatics*, vol. 4, no. 1, pp. 92-100(9), 2014.

[7] M. Chudecka, "Use of thermal imaging in the evaluation of body surface temperature in various physiological states in patients with different body compositions and varying levels of physical activity," *Central European Journal of Sport Sciences and Medicine*, vol. 1, no. 2, 2013.

[8] J. Wang, K. J. Chang, C. Y. Chen, K. L. Chien, Y. S. Tsai, Y. M. Wu, Y. C. Teng & T. T. F. Shih, "Evaluation of the diagnostic performance of infrared imaging of the breast: a preliminary study," *Journal of Biomedical Engineering Online*, vol. 9, no. 3, 2010.

[9] R. Freitas, *Nanomedicine* Volume I: Basic Capabilities. Landes Bioscience, Georgetown, USA, 1999.

[10] J. Kwok & J. Krzyspiak, "Thermal imaging and analysis for breast tumor detection," 2007.

[11] H. Qi, P. T. Kuruganti & W. E. Snyder, "Detecting breast cancer from thermal infrared images by asymmetry analysis," in M. Diakides and J. D. Bronzino & D. R. Peterson (Eds.), *Medical Infrared Imaging: Principles and Practice*, pp. 13.1-13.14, New York: Springer, 2012

[12] V. N. Vapnik, *The nature of statistical learning theory*, Springer- New York, NY, USA, 1995.

[13] Thirumuruganathan, S. (2010) Resource Document: A Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm. https://saravananthirumuruganathan.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighbor-knn-algorithm/.

[14] K. Cernak, "A comparison of decision tree classifiers for automatic diagnosis of speech recognition errors," *Computing and Informatics* vol. 29, no. 3, pp. 489-501, 2012.

[15] F. Amato, A. López, E. M. Peña-Méndez, P. Vaňhara, A. Hampl, & J. Havel, "Artificial neural networks in medical diagnosis," *Journal of applied biomedicine*, vol. 11, no. 2, pp. 47-58, 2013.