Middle-East Journal of Scientific Research 24 (Techniques and Algorithms in Emerging Technologies): 573-581, 2016
ISSN 1990-9233;

DOI: 10.5829/idosi.mejsr.2016.24.TAET23672

Breast Cancer Detection: A Framework to Classify Mammograms

[1]S. Devisuganya and [2]R.C. Suganthe

[1]Velalar College of Engineering and Technology, Erode, Tamilnadu, India
[2]Kongu Engineering College, Tamilnadu, India

Abstract: Breast cancer is a common life-threatening cancer affecting woman. Mammography is an effective screening tool; radiologists use for breast cancer detection. Mammogram, breast x-ray imaging is an effective, low cost, reliable method in early breast cancer detection. Mammogram images are classified as normal, benign and malignant. This study uses a feature selection method with Information gain to classify mammogram masses. The model's accuracy depends on relevant feature selection. Random Forests (RF) is successfully used for classification, but without information about classification reliability.

Key words: Pseudo Zernike Moments · Gaussian Markov Random Field · Information gain · C4.5 · Random Forest · Boosting techniques

INTRODUCTION

BREAST cancer is due to unrestrained growth of abnormal cells in the breast's milk-producing glands or in passages (ducts) that deliver milk to nipples. Breast cancer is a malignant (cancer) tumor starting with breast cells. It is found typically in women, but men to get breast cancer [1]. Breast cancer impacts the health care system that treats and monitors those diagnosed with the disease and provides end-of-life care for those who die from it [2].

There are many breast cancer types, but some are rare. These are Invasive Lobular Carcinoma (ILC), Ductal Carcinoma In Situ (DCIS), Invasive Ductal Carcinoma (IDC) and Inflammatory Breast Cancer (IBC) [3]. Screening is examining a group of people to detect disease or find those at increased disease risk [4]. Screening finds women with breast cancer to offer them early treatment [5]. Introduction of optimal screening should be followed by increase in rate of early disease succeeded by decrease in regional disease with the overall detection rate being constant [6]. Breast screening uses X-ray test called mammogram to check breasts for cancer signs [7].

Mammogram images are stored on film or a computer. A mammogram detects breast cancer early when it is small and when survival chances are the highest. Women above 40 years should have an annual mammogram [4]. It is possible for women to have breast cancer without it being visible on a mammogram. It is also possible for something that is not breast cancer to show up on a mammogram. This leads to additional tests, like diagnostic mammogram. Mammograms are better at locating breast cancer in older women than in those younger [8]. Mammography aims to detect, characterize and evaluate findings suggesting breast cancer and other breast diseases [9]. Screening and diagnostic are the 2 kinds of mammograms.

Screening mammograms are breast x-ray exams used for women who show no symptoms on breast or signs of breast cancer (like a prior abnormal mammogram). It involves two x-rays of each breast. Screening mammograms detect lumps and tumors that cannot be felt. They also find micro calcifications or tiny calcium deposits in the breast, which sometimes means that breast cancer is present. A screening mammogram aims to find breast cancer when it's too small to be felt by a woman or her doctor. Finding breast cancers early (before they grow and spread) greatly improves a woman's chance for successful treatment.

Diagnostic mammograms check for breast cancer after a lump or other symptoms or breast cancer signs are found. These include pain, thickened skin on breast, nipple discharge or a change in breast size and shape. This type of mammogram finds more about breast changes than on a screening mammogram, or views breast tissue that is tough to see on a screening mammogram [10]. A diagnostic mammogram is a breast X-ray examination when a patient shows signs and symptoms of breast disease, (abnormality detected on screening mammography) or who with earlier mammography findings needs an imaging follow-up [11].

Corresponding Author: S. Devisuganya, Velalar College of Engineering and Technology, Erode, Tamilnadu, India.

Two groups of women, identified based on individual attitude to screening, are compared to obtain an estimate between benefits and harms of mammography regarding absolute numbers of lives saved and absolute numbers of tumors over-diagnosed. To evaluate mammography screening programs benefits and harms, breast cancer mortality reduction and over-diagnosis was considered [12].

Regular mammogram benefits outweigh risks posed by small amount of radiation used [13]. Randomized controlled trials underestimate mammographic screening's true benefits. They evaluate mortality reduction among women randomized to a group invited to be screened instead of those who actually participated in screening. Breast cancer in younger women is more biologically aggressive [14]. So, mammography benefits depend on availability of effective treatment. Despite common misconceptions, screening mammography does not benefit women by reducing breast cancer risk, but by reducing mortality by detecting breast cancer at earlier and a more treatable stage [15].

Feature selection and feature extraction [16] are dimension reduction techniques. Feature selection is generally used in breast cancer classification. Feature selection filters redundant and irrelevant features from original data. Feature selection, a data mining pre-processing step selects and extracts valuable information in massive related materials. It is reported that logistic regression model discriminates between benign and malignant in decision making for early breast cancer detection and identifies most important features associated with breast cancer [17]. Feature selection's advantage including improvement of prediction reduces training times and ensures faster classifier performance [18].

This study proposes to use Information gain, C4.5, random tree and Boosting techniques. Section II deals with literature related to the work, Section III reveals methods used in the work, Section IV deals with results and discussions of obtained results and finally Section V concludes the work.

Literature Review: An unsupervised feature selection in mammogram image, using tolerance rough set based relative reduct was proposed by Aroquiaraj and Thangavel [19]. They compared it with Tolerance Quick Reduct and PSO - Relative Reduct unsupervised feature selection methods. A typical mammogram image processing system includes mammogram image acquisition, image segmentation pre-processing, feature extraction, feature selection and classification. The new method reduced features from extracted features and the method was compared to current unsupervised features selection methods. The new method was evaluated through clustering algorithms in K-means and WEKA.

A new unsupervised feature selection method using rough set based entropy measures was proposed by Thangavel and Velayutham [20]. A mammogram image processing system comprises image acquisition, pre-processing, segmentation, feature extraction, selection and classification. The new unsupervised feature selection method was compared to different supervised feature selection methods and evaluated with fuzzy c-means clustering to prove it efficiency in mammogram image classification.

An image classifier to classify mammogram images was proposed by Nugroho *et al*. [21]. The abnormality found in mammogram image was classified into malignant, benign and normal cases. Computer Aided Diagnosis (CAD) comprised 12 features of histogram and Gray-Level Co-occurrence Matrix (GLCM) as texture based features was extracted from mammogram image. Correlation based Feature Selection (CFS) reduced 50% of features. Multilayer perceptron algorithm was applied to mammography classification by selected features. Results showed that 40 digital mammograms data taken from private Oncology Clinic Kotabaru Yogyakarta achieved 91.66% accuracy.

A new fuzzy feature selection approach, which used fuzzy curve and fuzzy surface to select features from mammogram images, was introduced by Dubey *et al*. [22]. The approach used fuzzy curve to isolate a small set of major features from original features according to significance and eliminated unwanted features. Fuzzy surface eliminated features dependent on significant features reducing feature space dimensionality, thereby paving the way for a simplified classification scheme for practical applications. Results showed very promising features were chosen by this approach.

A technique to classify Regions Of Interests (ROIs) in digitized mammograms into mass and normal tissue regions proposed by Wong *et al*. [23] first found significant ROI texture features using Binary Particle Swarm Optimization (BPSO). Significant features are detected by a BPSO based feature selection technique. A decision tree classifier classified the test set and used significant features. Results showed that significant texture features located by BPSO based feature selection technique had better classification accuracy compared to a full features set.

Using random feature selection method for mammogram images classification using a multi-scale transform was discussed by Faye [24]. Every image was represented by a coefficients vector. Columns subsets were randomly generated and used for training set classification. Subsets achieving a predefined performance are pooled in a final set for testing. The method was tested with images set provided by Mammography Image Analysis Society (MIAS) to differentiate between normal and abnormal images. Classifiers K nearest neighbors and Discriminant Analysis (DA) were used with Wavelet transform in the experiments.

A new feature extraction method based on spectral shape for abnormality classification in mammogram images was proposed by Velayutham and Thangavel [25]. Spectral shape features were extracted from mammogram images and analyzed for classification performance. The method's classification performance was compared with Haralick features and run-length features. The processes were executed and features analyzed. The proposed spectral shape feature's performance was examined.

A CAD system for detection of normal or abnormal pattern in the breast was presented by Radovic *et al*. [26]. The new system had 4 steps: image pre-processing, feature extraction, feature selection and classification to classify mammogram images into normal (without tumor) and abnormal (with tumor) patterns. After noise was removed the mammogram used Discrete Wavelet Transformation (DWT), which chose the ROI. A total of 20 GLCM features were extracted from ROI and were inputs for the classification algorithms.

A new automatic breast abnormality detection method that used mammogram images to determine breast tissue abnormality was introduced by Lashkari [27]. Gabor wavelets, Geometric Moment Invariants, energy, entropy, contrast and statistic features like mean, median, variance, correlation, values of maximum and minimum intensity were used to provide a clear description from breast tissue. It uses feature selection to reduce feature space. This project aims to classify breast tissues into normal and abnormal classes automatically, saving radiologist's time and increasing accuracy.

A new lesion detection algorithm based on layer structuring hypothesis where different layers are obtained with different thresholds adaptively determined from mammogram histogram was presented by Zhou and Wang [28]. Highly suspicious lesion regions were obtained from selection procedures based on morphological features and Single Concentric Layers Criterion. 170 mammograms were selected from MIAS dataset for the proposed algorithms evaluation. Results indicated that the method had potential to aid radiologists in mammogram interpretations.

A Classification method for normal and abnormal tissues in mammograms using curvelet transform was presented by Eltoukhy *et al*. [29]. Curvelet coefficients were represented in certain coefficients groups independently. Statistical features were calculated for every group of coefficients. These were combined with features from mammogram image itself. To improve classification rate, feature ranking was applied to select most significant features. SVM classification results used 10-fold cross validation and are presented here. Classification results showed that ranked features improved classification rate by up to 85.48% with a group of 200 coefficients.

MATERIALS AND METHODS

Features are observable image patterns which provide information of the image. Classification accuracy depends on feature extraction. Feature denotes a piece of information relevant to solving a computational task related to a specific application. Specifically, features refer to result of a general neighborhood operation applied to an image, specific structures in image itself, ranging from simple structures like points/edges to more complex structures like objects. Many features were extracted for mammogram abnormalities. Texture feature extraction methods play an important role in detecting mammogram abnormalities due to its nature [30]. This section discusses C4.5, random forest and Boosting techniques. Pseudo Zernike Moments and Gaussian Markov Random Field (GMRF) were used for feature extraction.

A Pseudo Zernike Moments: The Zernike moments computation of an input image has 3 steps - computation of 1) radial polynomials, 2) Zernike basis functions and 3) Zernike moments by projecting image onto Zernike basis functions [31].

The kernel of pseudo-Zernike moments is orthogonal pseudo-Zernike polynomials set defined over polar coordinate space in a unit circle. The 2-dimensional pseudo-Zernike moments of order p with repetition q of an image intensity function is defined as [32]:

$$Z_{pq} = \frac{p+1}{\pi}\int_{-\pi}^{\pi}\int_{0}^{1} V_{pq}^{*}(r,\theta) f(r,\theta) r dr d\theta; \quad |r| \le 1 \tag{1}$$

where pseudo-Zernike polynomials V_{pq} of order p are defined as:

$$V_{pq}(r,\theta) = R_{pq}(r)e^{\hat{j}q\theta}; \qquad \hat{j} = \sqrt{-1} \tag{2}$$

and the real-valued radial polynomials, $R_{pq}(r)$, is given as:

$$R_{pq}(r) = \sum_{k=0}^{p-|q|}(-1)^k \frac{(2p+1-k)!}{k!(p+|q|+1-k)!(p-|q|-k)!} r^{p-k} \tag{3}$$

where $0 \le |q| \le p$.

As pseudo-Zernike moments are defined regarding polar coordinates (r, θ) with $|r|\le 1$, computation of pseudo-Zernike polynomials requires a linear transformation of image coordinates (i, j), $i, j = 0, 1, 2, \ldots, N\text{-}1$ to a suitable domain $(x, y) \in R^2$ inside a unit circle. Two commonly used cases of transformations. Based on these, following discrete approximation of continuous pseudo-Zernike moments' integral.

$$Z_{pq} = \lambda(p,N)\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} R_{pq}(r_{ij})e^{-\hat{j}q\theta_{ij}} f(i,j), \quad 0 \le r_{ij} \le 1 \tag{4}$$

where most general image coordinate transformation to interior of unit circle is given by;

$$r_{ij} = \sqrt{(c_1 i + c_2)^2 + (c_1 j + c_2)^2}, \ \theta_{ij} = \tan^{-1}\left(\frac{c_1 j + c_2}{c_1 i + c_2}\right) \tag{5}$$

Gaussian Markov Random Field (GMRF): Let $x = (x_1, x_2,\ldots,x_n)^T$ be a Gaussian random field with mean μ and covariance matrix Σ, that is, $x \sim N(\mu, \Sigma)$. The precision matrix of x is denoted by Q and $Q = \Sigma^{-1}$. Gaussian random field x is said to be a Gaussian Markov Random Field (GMRF) regarding labeled undirected graph $G = (V, \varepsilon)$, if nodes are $V = \{1, \ldots, n\}$ and edges;

$$\varepsilon = \{\{i,j\} \in V \times V : Q_{ij} \ne 0 \ and \ i \ne j\}. \tag{6}$$

If $\{i, j\} \in \varepsilon$, then i and j are said to be neighbors and is written as $i\text{\textasciitilde}j$. Further, notation $x_{i:j}$ to refer to sub-vector of x corresponding to nodes $i, i\text{+}1, \ldots, j$. By definition, any GRF is a GMRF, generally regarding a fully connected graph G.

In practice, use of GMRFs is confined to situations where neighborhood size is small so that precision matrix is sparse. The precision matrix's non-zero pattern is related to conditional independence structure of GMRF by $x_i \perp x_j \mid x_{-ij} \Leftrightarrow Q_{ij} = 0, i \ne j$.

Here, x_{-ij} denotes all elements of x except elements i and j. As a consequence of correspondence between non-zero pattern of Q and conditional independence structure of GMRF, GMRF is specified regarding its conditional moments [33].

A mammographic image Y is modeled by a finite lattice GMRF. Each pixel in image lattice L is represented by a random variable y_{ij} where $Y = \{y_{ij}: 0\le i \le M-1, 0 \le i \le M-1\}$ and $L = \{(i, j): 0\le i \le M-1, 0 \le i \le M-1\}$. In a GMRF assumption of image Y with respect to a certain neighborhood system η, Y is reshaped to a single vector $y = [y_1, y_2, \ldots\ldots\ldots, y_M{}^2$ in lexicographic order [34].

Normalization: The process where the range of pixel intensity values changes is called Normalization. Applications, for example, include photographs with poor contrast due to glare. Normalization is also called contrast stretching. In data processing fields like digital signal processing, it is called dynamic range expansion whose purpose are various applications it brings to the image, or other type of signal, to a range that is familiar or normal to senses. So it is called normalization. Feature pre-processing "enhances" features for classification. Features are transformed to normalize scatter of distribution and enhance separation distance between 2 classes. A well-known transform is "whitening" transform used to make transformed features "in-dependent" [35] [36]. The extracted Pseudo Zernike Moments and GMRF features are normalized and concatenated into a single feature subset after normalization.

Information Gain (IG): Information Gain (IG) measurement normalized with symmetrical uncertainty coefficient is a symmetrical measure where amount of information gained about Y after observing X is equal to amount of information gained about X after observing Y (a measure of feature inter-correlation) [37]. Information gain is based on entropy decrease after a dataset is split on an attribute.

$$\text{Information Gain (S, A)} = \text{Entropy(S)-H(S, A)} \tag{7}$$

where $H(S, A) = \Sigma_i (|S_i|/|S|).H(S_i)$

A takes on value 1 and $H(S_i)$ is entropy of system of subsets Si.

Training data is a set $S=s_1, s_2$—of already classified samples based on mathematical morphological and new features. Each sample $S_i = x_1, x_2$ is a vector where x_1, x_2—represents attributes or features of sample [38].

Classifier: Classification is divided into training phase and testing phase. In training phase, known data is given and features calculated by processing, which precedes classification [39]. The classifier then classifies whether an entire whole-field mammogram is normal. But, in such binary tree classifiers, errors may accumulate from one level to another, making classification erroneous. Random Forests Decision Classifier (RFDC), involving regression trees, was used in mammogram classification.

Random forests classifier is used for masses classification with geometry and texture features [40]. But, a major problem is the large number of features which find it hard to determine which feature or features combination achieves better classification accuracy. So, it is important to select a suitable and optimized features set from a high dimensional feature matrix with the ability to differentiate between different mammogram types [41].

A C4.5

Entropy and information gain are used for tree splitting by C4.5 classification. It handles categorical and continuous data. A threshold value is fixed so that all values above a threshold are not considered. The initial step is calculating information gain for every attribute. An attribute with maximum gain is preferred as decision tree root node.

C4.5, in a set S of cases, first grows an initial tree using divide-and-conquer algorithm as:

- If all cases in S belong to same class or S is small, tree is a leaf labeled with most frequent class in S.
- Otherwise, choose a test based on one attribute with 2 or more outcomes. Make test root of tree with one branch for every outcome of test, partition S into corresponding subsets S1, S2, ...according to outcome for every case and apply same procedure recursively to every subset [42].

This problem is addressed by C4.5 (decision tree) along with 2 sampling techniques, which handle class imbalance and contribute to masses classification. C4.5 is the classifier with above sampling techniques as it was used in imbalanced domains [43].

Typically, C4.5 assigns frequency of correct counts at leaf as probabilistic estimate. For notational purposes, TP is number of true positives at leaf; FP is number of false positives and C number of classes in data set. Thus, frequency based probabilistic estimate is written as [44]:

$$\text{Pleaf} = \text{TP} = (\text{TP} + \text{FP}) \quad (8)$$

Random Forest: Random Forest (RF) classifiers compared favorably with common classifiers. RF is an ensemble learning technique, which combines many decision trees to make a prediction, giving as output a class that is a mode of classes output by individual trees. RFs are a family of methods, of different decision tree ensemble induction algorithms like Breiman Forest-RI method where training set for every individual tree in a Random forests is constructed by sampling N examples randomly with replacement from N available examples in a dataset. This is bootstrap sampling and bagging describes aggregation of predictions from resulting trees collection. "Out-of-bag" predictions are those derived from non-bootstrapped observations, which built a particular tree.

Forest-RI Algorithm grows a decision tree using the process:

Let T be number of trees to build, for every |T| iteration.

- Select a new bootstrap sample from training set.
- Grow un-pruned tree on bootstrap.
- At each internal node, arbitrarily select m predictors and determine best split using only the predictors.
- Output overall prediction as majority vote from all individually trained trees [40].

RF also has 2 built-in heuristics to estimate variable importance, allowing insight into data structure and is very robust against over fitting. It is easy to optimize, as performance depends on 2 parameters.

Random forest algorithm is:

For b = 1;….; T: Create a bootstrap sample *Lb* by randomly drawing N samples with replacement from N samples in learning set *L*. Use L_b to build a tree:

- At node n, randomly sample m of M predictor variables.
- For each of m sampled variables v_k, k = 1;…; m find best split s_k among all possible splits.
- Choose best split s^* among k = 1;….; m splits s_k on which to split a node. This variable v_{best} at is identified cut point c^* splits node.
- Split all data entries i = 1;….; n, in parent node, by sending observations with $v^i_{best} < c^*$ to left descendant node and all observations $v^i_{best} \geq c^*$ to right descendant node [45].

- Repeat steps 1-4 on all descendant nodes to grow a maximally sized tree T_h.

Boosting Techniques: Boosting algorithms are based on the ideas that sum of weak classifiers produces a strong classifier. In Gentleboost algorithm, weak classifiers (*ht*) are simple regression stumps with one of features, so at each round *t* feature with less error is chosen. A weak classifier used is [46]:

$$h_t(x) = a\delta(x_i > th) + b \tag{9}$$

where *th* is a threshold determining if pattern *x* belongs to object class, *xi* is *i*th dimension of *x* and *a* and *b* are parameters selected to minimize classifier error (*a* is regression slope and *b* offset):

$$e = \sum\left(z\left(y - \left(a\left(x_i > th\right) + b\right)\right)^2 \right) \tag{10}$$

At each round training data weights (*z*) are updated, increasing in following round a possibility of classifying correctly earlier incorrectly classified points. In GentleBoost algorithm data weights are updated:

$$z_{t+1} = z_t e^{y.\mathrm{h}_t(x)} \tag{11}$$

So, when testing new data, final (strong) classifier is computed using weak classifier created at every round of boosting. So, testing data is classified according to sign of a sum of weak classifiers:

$$H(x) = \Sigma h(x) \tag{12}$$

Absolute value of *H*(*x*) shows classified data's confidence [47].

RESULT

To evaluate the various techniques, 150 normal mammogram image and 25 image with calcification obtained from MIAS dataset were used. Features are extracted using Pseudo Zernike Moments and Gaussian Markov Random Field technique. The extracted features are concatenated after normalization. Best Features are selected using IG. Classification is achieved using C4.5, random tree and Boosting techniques. Results are presented in this section.

It can be observed from Fig. 1, IG with C4.5 improved the classification accuracy by 5.59%. Similarly, IG with random forest increased classification accuracy by 3.44%.

Table 1: Classification Accuracy

	Classification accuracy %
C4.5	79.43
Random Forest	81.71
Boosting	82.86
IG and C4.5	84
IG and Random Forest	84.57
IG and Boosting	85.14

Table 2: Sensitivity

	Sensitivity
C4.5	0.76
Random Forest	0.76
Boosting	0.76
IG and C4.5	0.76
IG and Random Forest	0.8
IG and Boosting	0.8

Table 3: Specificity

	Specificity
C4.5	0.8
Random Forest	0.8267
Boosting	0.84
IG and C4.5	0.8533
IG and Random Forest	0.8533
IG and Boosting	0.86

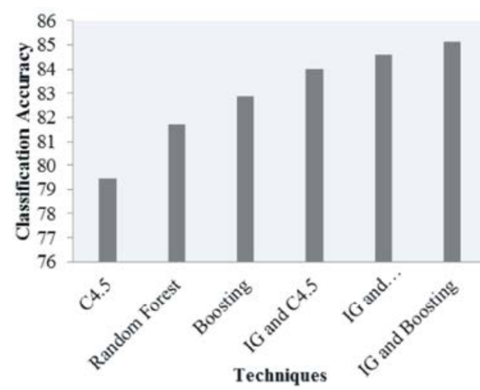

Fig. 1 Classification Accuracy

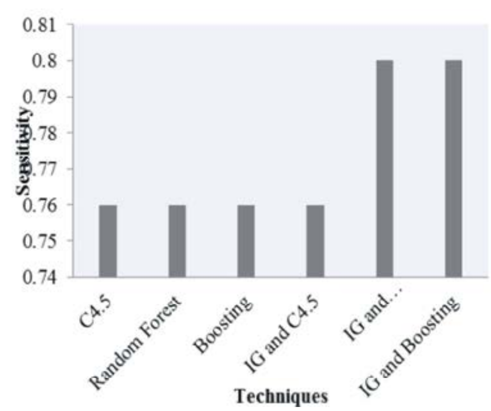

Fig. 2: Sensitivity

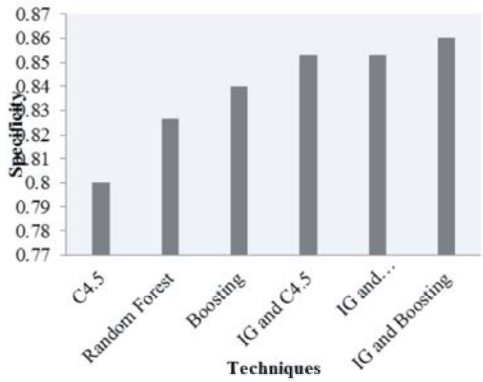

Fig. 3: Specificity

It can be observed from Fig. 2, IG with boosting improved the sensitivity by 5.13%. Similarly, IG with random forest improved the sensitivity by 5.13%.

It can be observed from Fig. 3, IG with random forest improved the specificity by 3.17%. Similarly, IG with C4.5 improved the specificity by 6.45%.

CONCLUSION

Mammography is one of the best breast cancer detection methods. But, in some cases, radiologists face problems in detecting tumours. Methods like the one presented in this paper could help medical staff to improve detection accuracy. Early diagnosis through regular screening and timely treatment prevents cancer. This study presented a new approach to segment breast cancer mass in mammograms. The study focuses on improving classification performance through feature selection. It is seen that of the various classification techniques C4.5 outperforms other algorithms with highest accuracy.

REFERENCES

1. Moynihan, R., I. Heath and D. Henry, 2002. Selling sickness: the pharmaceutical industry and disease mongeringCommentary: Medicalisation of risk factors. Bmj, 324(7342): 886-891.
2. Bergman, A., J.J. Heindel, T. Kasten, K.A. Kidd, S. Jobling, M. Neira and T.J. Woodruff, 2013. The impact of endocrine disruption a consensus statement on the state of the science. Environmental Health Perspectives, 121(4): A104.
3. Early Breast Cancer Trialists' Collaborative Group, 1998. Tamoxifen for early breast cancer: an overview of the randomised trials. The Lancet, 351(9114): 1451-1467.
4. Baker, L.H., 1982. Breast cancer detection demonstration project: Five-year summary report. CA: a Cancer Journal for Clinicians, 32(4): 194-225.
5. Gøtzsche, P.C. and M. Nielsen, 2009. Screening for breast cancer with mammography. The cochrane library.
6. Esserman, L., Y. Shieh and I. Thompson, 2009. Rethinking screening for breast cancer and prostate cancer. Jama, 302(15): 1685-1692.
7. Gøtzsche, P.C., O.J. Hartling, M. Nielsen, J. Brodersen and K.J. Jørgensen, 2009. Breast screening: the facts-or maybe not. BMJ, pp: 338.
8. Kreuter, M.W., K. Holmes, K. Alcaraz, B. Kalesan, S. Rath, M. Richert and E.M. Clark, 2010. Comparing narrative and informational videos to increase mammography in low-income African American women. Patient Education and Counseling, 81: S6-S14.
9. Sickles, E.A., D.E. Wolverton and K.E. Dee, 2002. Performance parameters for screening and diagnostic mammography: specialist and general radiologists 1. Radiology, 224(3): 861-869.
10. Elmore, J.G., M.B. Barton, V.M. Moceri, S. Polk, P.J. Arena and S.W. Fletcher, 1998. Ten-year risk of false positive screening mammograms and clinical breast examinations. New England Journal of Medicine, 338(16): 1089-1096.
11. Lee, C.H., D.D. Dershaw, D. Kopans, P. Evans, B. Monsees, D. Monticciolo and L.W. Burhenne, 2010. Breast cancer screening with imaging: recommendations from the Society of Breast Imaging and the ACR on the use of mammography, breast MRI, breast ultrasound and other technologies for the detection of clinically occult breast cancer. Journal of the American College of Radiology, 7(1): 18-27.
12. Puliti, D., G. Miccinesi, M. Zappa, G. Manneschi, E. Crocetti and E. Paci, 2012. Balancing harms and benefits of service mammography screening programs: a cohort study. Breast Cancer Res., 14(1): R9.
13. Fletcher, S.W. and J.G. Elmore, 2003. Mammographic screening for breast cancer. New England Journal of Medicine, 348(17): 1672-1680.
14. Berg, W.A., 2010. Benefits of screening mammography. Jama, 303(2): 168-169.
15. Rimer, B.K., P.A. Briss, P.K. Zeller, E.C. Chan and S.H. Woolf, 2004. Informed decision making: what is its role in cancer screening?. Cancer, 101(S5): 1214-1228.

16. Surendiran, B. and A. Vadivel, 2010. Feature selection using stepwise anova discriminant analysis for mammogram mass classification. International J. Recent Trends in Engineering and Technology, 3(2): 55-57.
17. Luo, S.T. and B.W. Cheng, 2012. Diagnosing breast masses in digital mammography using feature selection and ensemble methods. Journal of Medical Systems, 36(2): 569-577.
18. Nithya, R. and B. Santhi, 2011. Mammogram classification using maximum difference feature selection method. Journal of Theoretical and Applied Information Technology, 33(2).
19. Aroquiaraj, I.L. and K. Thangavel, 2013. Mammogram image feature selection using unsupervised tolerance rough set relative reduct algorithm. InPattern Recognition, Informatics and Mobile Engineering (PRIME), International Conference on, pp: 479-484). IEEE.
20. Thangavel, K. and C. Velayutham, 2012. Rough set based unsupervised feature selection in digital mammogram image using entropy measure. In Biomedical Engineering (ICoBE), International Conference on (pp: 10-16). IEEE.
21. Nugroho, H.A., N. Faisal, I. Soesanti and L. Choridah, 2014. Identification of malignant masses on digital mammogram images based on texture feature and correlation based feature selection. In Information Technology and Electrical Engineering (ICITEE), 6th International Conference on (pp: 1-6). IEEE.
22. Dubey, R.B., M. Hanmandlu and S. Vasikarla, 2014. Features Selection in Mammograms Using Fuzzy Approach. In Information Technology: New Generations (ITNG), 2014 11th International Conference on (pp: 123-128). IEEE.
23. Wong, M.T., X. He, H. Nguyen and W.C. Yeh, 2012. Faculty of Engineering & Information Technology, University of Technology, Sydney, Broadway, NSW 2007, Australia. In Computerized Healthcare (ICCH), 2012 International Conference on (pp: 152-157). IEEE.
24. Faye, I., 2012. A random feature selection method for classification of Mammogram images. In Intelligent Systems, Modelling and Simulation (ISMS), 2012 Third International Conference on (pp: 330-333). IEEE.
25. Velayutham, C. and K. Thangavel, 2011. A novel feature extraction method using spectral shape in digital mammogram image. In Information and Communication Technologies (WICT), World Congress on (pp: 835-840). IEEE.
26. Radovic, M., M. Djokovic, A. Peulic and N. Filipovic, 2013. Application of data mining algorithms for mammogram classification. InBioinformatics and Bioengineering (BIBE), IEEE 13th International Conference on (pp: 1-4). IEEE.
27. Lashkari, A., 2010. Full automatic micro calcification detection in mammogram images using artificial neural network and Gabor wavelets. InMachine Vision and Image Processing (MVIP), 6th Iranian (pp: 1-7). IEEE.
28. Zhou, Y. and J. Wang, 2013. Mammographic lesion detection based on improved concentric morphology model. In Mechatronic Sciences, Electric Engineering and Computer (MEC), Proceedings 2013 International Conference on (pp: 929-932). IEEE.
29. Eltoukhy, M.M., S. Gardezi, S. Jamal and I. Faye, 2014. A method to reduce curvelet coefficients for mammogram classification. In Region 10 Symposium, IEEE (pp: 663-666). IEEE.
30. Pradeep, N., H. Girisha, B. Sreepathi and K. Karibasappa, 2012. Feature extraction of mammograms. International Journal of Bioinformatics Research, ISSN, 0975-3087.
31. Tahmasbi, A., F. Saki and S.B. Shokouhi, 2011. Classification of benign and malignant masses based on Zernike moments. Computers in Biology and Medicine, 41(8): 726-735.
32. Chong, C.W., P. Raveendran and R. Mukundan, 2003. The scale invariants of pseudo-Zernike moments. Pattern Analysis & Applications, 6(3): 176-184.
33. Rue, H. and T. Follestad, 2004. Efficient Computations for Gaussian Markov Random Field Models with two Applications in Spatial Epidemiology.
34. Chen, C.H. and G.G. Lee, 1997. On digital mammogram segmentation and microcalcification detection using multiresolution wavelet analysis. Graphical Models and Image Processing, 59(5): 349-364.
35. Omran, M., A.P. Engelbrecht and A. Salman, 2005. Particle swarm optimization method for image clustering. International Journal of Pattern Recognition and Artificial Intelligence, 19(03): 297-321.
36. Sun, Y., 2004. Normal mammogram analysis (Doctoral dissertation, Purdue University).
37. Pérez, N., M.A. Guevara, A. Silva, L. Ramos and J. Loureiro, 2014. Improving the performance of machine learning classifiers for Breast Cancer diagnosis based on feature selection. In Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on (pp: 209-217). IEEE.

38. Angayarkanni, S.P., N.A.D.I.R.A. Banu Kamal and V. Thavavel, 2013. Mammogram Mri Image Segmentation And Classification. Journal Of Theoretical & Applied Information Technology, 57(2).
39. Qian, W., L.P. Clarke, B. Zheng, M. Kallergi and R. Clark, 1995. Computer assisted diagnosis for digital mammography. Engineering in Medicine and Biology Magazine, IEEE, 14(5): 561-569.
40. Shili, H., L.B. Romdhane and B. Ayeb, 2013. Reliable Probabilistic Classification of Mammographic Masses using Random Forests. In The 9th International Conference on Data Mining (DMIN'2013)(Las Vegas, Nevada, USA).
41. Beura, S., B. Majhi and R. Dash, 2014. Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer. Neurocomputing.
42. Subasini, A. and N.F. Abubacker, 2014. Analysis of classifier to improve Medical diagnosis for Breast Cancer Detection using Data Mining Techniques. International Journal of Advanced Networking & Applications, 5(6).
43. Murthy, S.N., A. Kumar and H.S. Sheshadri, Abnormality Detection in Digital Mammogram.
44. Mohanty, A.K. and S.K. Lenka, Efficient image mining technique for classification of mammograms to detect breast cancer. Special Issue of IJCCT, 2(2): 3.
45. Baumgartner, C., 2010. Computer-Aided Detection of Malignant Breast Masses in Mammography using Random Forests.
46. Oliver, A., A. Torrent, M. Tortajada, X. Lladó, M. Peracaula, L. Tortajada and J. Freixenet, 2010. A boosting based approach for automatic micro-calcification detection. In Digital Mammography (pp: 251-258). Springer Berlin Heidelberg.
47. Torrent, A., A. Oliver, X. Lladó, R. Marti and J. Freixenet, 2010. A supervised micro-calcification detection approach in digitised mammograms. In ICIP, (pp: 4345-4348).