Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 43 No. 7
July 2023

A Robust Technique Using Pruned Association Rule to Diagnose Breast Cancer from Mammograms

Manmohan Shoo
KMBBCET, Khorda
for_maumohan@yahoo.co.in

Amalendu Bag
RCM College,BBSR
amalendu.bag@gmail.com

Aswini Kumar Mohanty
The Tecno School, BBSR
asw_moh@yahoo.com

Abstract-Many researches have been carried out to diagnose early breast cancer and many authors have suggested many methodologies but have not been widely accepted either by the radiologist or oncologist so far due to many reasons. A robust technique for breast tumor classification using pruned association rule PARM algorithm is presented in this paper believed to be acceptable by the physician. The method proposed makes use of association rule mining technique to classify the mammogram breast images into two categories namely benign and malign. It combines the GLCM features extracted from images and high level knowledge from specialists. The developed algorithm can assist the Radiologist for effectively classify with multiple keywords per image to improve the accuracy. Here using Association rule mining performs benign-malignant classification on region of interest (ROI) that contains mass. Texture is one of the important characteristic for features to classify both from film reading and machine learning. ARM exploits this important factor to classify the mass into benign or malignant. The statistical textural features used in characterizing the masses are mean, standard deviation, entropy, skewness, kurtosis and uniformity which are standard features for classification purpose to avoid under fitting or over fitting for the classifier. The main aim of the method is to increase the effectiveness and efficiency of the classification process with an objective to reduce the numbers of false-positive and increasing the sensitivity of malignancies. Association rule mining was proposed for classifying the marked regions into malignant as we as benign are 92.30% sensitivity and 95.23% specificity which is very much encouraging in compare to the radiologist's sensitivity 80%.

Keywords-Pruned Association rule mining, Digitized Mammograms, Texture Features, Gray Level Co-occurrence Matrix, The Apriori Algorithm

1. Introduction

Breast cancer is a public health problem in the world. Worldwide, it is the leading cause of death for women in their 40's [1] and second after lung cancer. In one European country in 2019, an estimated 42,600 women and 170 men were diagnosed with breast cancer and 11,400 women and 67 men died from it. Therefore, 1 in 10 women (10%) is expected to develop breast cancer during her lifetime (by age 70) and 1 in 25 will die from it. Although the breast cancer morbidity rates have increased over the years, breast cancer mortality has declined among women of all ages [1]. This positive trend in mortality reduction may be associated with improvements made in early detection of breast cancer, treatment at an earlier stage and the broad adoption of x-ray mammography [1]. However, there still remain significant sides for improvements to be made in x-ray mammography since they are primarily based on the ability of expert radiologists in detecting abnormalities.

Mammography has been one of the most reliable methods for early detection of breast carcinomas. X-ray mammography is currently considered as standard procedure for breast cancer diagnosis. However, retrospective studies have shown that radiologists can miss the detection of a significant proportion of abnormalities in addition to having high rates of false positives. The estimated sensitivity of radiologists in breast cancer screening is maximum about 80% [2]. Double reading has been suggested to be an effective approach to improve the sensitivity. But it becomes costly because it requires twice as many radiologists' reading time. This cost will be quite over burden considering the ongoing efforts to reduce costs of the health care system.

The bigger ambiguity is one of the major requirements for a mass screening program to be successful. The ultimate diagnosis of all types of breast disease depends on a biopsy. In most cases the decision for a biopsy is based on mammography findings. Biopsy results indicate that 65-90% of suspected cancer detected by mammography turned out to be benign [3]. Therefore, it would be valuable to develop a computer aided method for mass classification based on extracted features from the region of interests (ROI) in mammograms. This would reduce the unnecessary biopsies in patients with benign disease and thus avoid patients' psychological suffering, with an added bonus of reducing healthcare costs. The principal stages of computer-aided breast cancer detection and classification is shown in figure 1.Tthe ROIs extracted from the breast region are shown in figure 2 and masses extracted from the ROIs both for malignant and benign are shown in figure 3 that need to be classified and it is the main theme of the paper. In this paper automatic mass classification into benign and malignant is presented based on the statistical and textural features extracted from mass from the breast region using proposed ARM. This paper is organized as follows. Section 2 briefly reviews some existing techniques for mass classification followed by Association rule mining with correlation in section 3. Statistical texture features are described in section 4. Section 5 describes the materials and proposed methods for mass classification. Section 6 demonstrates some simulation results and their performance evaluation finally conclusions are presented in section 7.

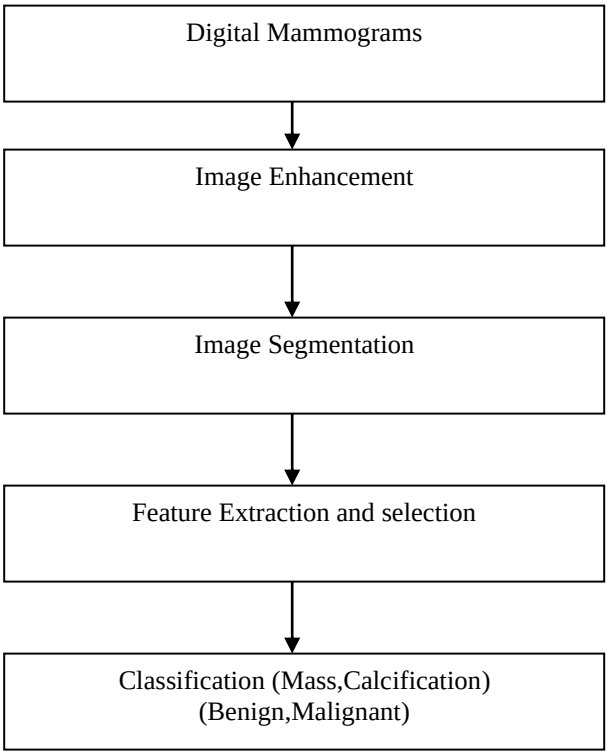

Figure 1: Sample mammogram and principal stages of breast cancer detection

2. Literature survey

Masses are one of the important early signs of breast cancer. They are often indistinguishable from the surrounding parenchyma because their features can be obscured or similar to the normal inhomogeneous breast tissues. The gray levels of those inhomogeneous tissues in the breast could vary with the distribution of breast soft tissue. Furthermore, the difficulty could be increased due to the fact that the masses in digitized mammograms are similar to the glands, cysts or dense portion of the breast [4]. This makes the automatic mass detection, segmentation and classification challenging. The main aim of this paper is to develop a classifier for breast cancer detection of masses in mammograms.

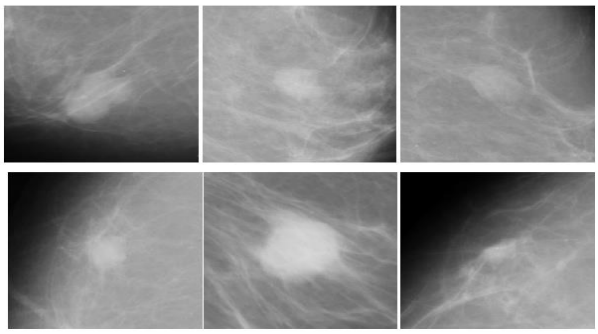

Figure 2: Sample ROIs extracted from breast region

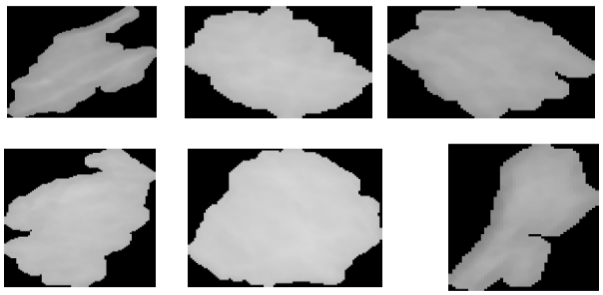

Figure 3: Sample masses extracted from the ROIs

In recent years, a few researchers have used different approaches to do the classification of masses. The classification step is vital for the performance of the computer aided diagnosis (CAD) system that is shown in figure 1. Many methods involve the use of classification techniques to classify the marked region in the mammogram. Classifiers such as linear discriminants analysis (LDA) and ANN have performed well in mass classification. Other classifiers Bayesian network and binary decision tree and support vector machine (SVM) are also used in this application.

LDA's [3, 5, 6] are traditional method for classification. They construct decision boundaries by optimizing certain criteria to classify cases into one of mutually exclusive classes. They show high performance for linear separable problems but poor for nonlinear separable data. LDA in combination with stepwise feature selection [3, 7] was trained and tested on morphological features extracted using the machine segmentation and radiologist segmentation and Az area under the receiver operating characteristics (ROC) 0.89 was btained whereas for speculation features it was 0.88.

Bayesian network uses a probabilistic approach to determine the class conditional probability density functions for background and tumor in breast cancer detection application. In [8] two phases hierarchical scheme is used to classify the masses where Bayesian classifier exists in each level. In first phase the speculated masses are discriminated from non-speculated masses. In the second phase masses with fuzzy edges are separated from well defined edges among the non-speculated edges.

ANNs are invaluable tools in various medical diagnostic systems. Lisboa [9] reviewed the improvements in health care arising from the participation of NN in medical field. ANN constructively makes use of trained data set to make complex decisions. It is robust, no rule or explicit expression is needed and widely applicable [3]. But there is no common rule to determine the size of the ANNs, long training time and sometimes over training. Alginahi et. al. [10, 11] developed ANN-based technique for thresholding composite digitized documents with non-uniform and complex background. It was used in the application of segmenting bank cheques from complex background for application in OCR. This method uses 8 statistical and textural features of an image.

A common database and the same genetic algorithm were used to optimize both the Bayesian belief network and neural network in [3, 12, 13]. The results show that the performance of the two classifiers converged to the same level. Therefore, it is obvious that the performance of CAD systems mostly depends on feature selection and training database than the classifiers.

Antonie et al [14] use association rules to classify digital mammograms into normal, benign, and malignant classes. However, their technique is time consuming, requires labeling of quadrants with abnormalities, and relies on very low support and confidence values, resulting in the generation of weak rules. Ribeiro et al. [15] use texture features and association rules to classify mammogram images. The major problems with

this technique are the ad hoc segmentation of images, the time consuming discretization of segments, and the constraint of keeping the class label on right side of the rule. In [16], Yun et al. use a combination of association rules with a rough set theory for mammogram classification.

In addition to the above, Tseng et al. [17] apply multilevel association rules to hierarchically clustered objects from various images and perform object based segmentation on the image. This technique is not widely applicable for medical images because they usually contain few objects and because the objects may contain different abnormalities, i.e. different stages of cancer in breast mammograms.

In this paper, we present a new association rule based technique for mammogram image classification. We extract texture features from images to form association rules, which are then employed for classifier building and validation. Rigorous experimentation is performed, and we achieve superior classification accuracy on a previously studied mammogram dataset, demonstrating the efficacy of our technique.

3. Association Rule Mining

Association rule mining is a well-known technique in data mining. It is able to reveal all interesting relationships, called associations, in a potentially large database. However, how interesting a rule is depends on the problem a user wants to solve. Existing approaches employ different parameters to guide the search for interesting rules. Classification using association rules combines association rule mining and classification, and is therefore concerned with finding rules that accurately predict a single target (class) variable.

The key strength of association rule mining is that all interesting rules are found. The number of associations present in even moderate sized databases can be, however, very large usually too large to be applied directly for classification purposes.

Association Rule are the out put of the process of finding Associations or correlations among an item set. The rule form is LHS ->RHS [support, Confidence]

3.1 Classification using association rules

Classification using association rules can be divided into three fundamental parts:

1. Association rule mining,
2. Pruning and
3. Classification.

The mining of association rules is a typical data mining task that works in an unsupervised manner. A major advantage of association rules is that they are theoretically capable of revealing all interesting relationships in a database. But for practical applications the number of mined rules is usually too large to be exploited entirely. This is why the pruning phase is stringent in order to build accurate and compact classifiers. The smaller the number of rules a classifier needs to approximate the target concept satisfactorily, the more human-interpretable is the result.

3.2 Problem Definition

Let database D is a set of instances where each instance is represented by $< a_1, a_{2.....} a_m, c>$, where $a_1, a_{2.....} a_m$ are non class attributes and c is a class attribute C. A common rule is defined as $x \rightarrow c$,

Where x is a set of non class attributes and c is class label.

A rule classifies an instance if the instance contains all the attribute values in X and the class label c.

The quality measurement factor of a rule is support and confidence where support is (num(x, y) / |D|), |D| denotes the total number of instances in database and confidence is (num(x, y) / num(x)). Rule items that satisfy min_sup are called frequent rule items, while the rest are called infrequent rule items.

The task is to generate the CARs that satisfies both min_sup and min_conf constraints. These CARs can be used to build a classifier which would be able to classify new instance accurately. The input to build a classifier is a pre processed set of Class Association Rules.

3.3. Association Rule Mining for Classification

Algorithm: PARM Find association on the training set of the transactional database. +
Input: A set of Image patches (P1) of the form P_i:{k_1,k_2,……k_m,f_1,f_2,……f_n}
where k_i is a keyword attached to the patches and f_j are the selected features for the patches, a minimum support threshold σ .
Output: A set of association rules of the form
$f_1 \wedge f_2 \wedge$.....$\wedge f_n \Rightarrow k_i$
where k_i is a keyword, f_j is a feature and kw is a class category.

Method:
(1) ← C_0 {candidate keywords and their support}
(2) ← F_0 {frequent keywords and their support}
(3) ← C_1 {candidate keyword 1 item sets and their support}
(4) ← F_1 {frequent 1 item sets and their support}
(5) ← C_2 {candidate pairs (k, f), such that (k, f)∈P_1 and k∈F_0 and f ∈F_1}
(6) For each patches p in P_1 do {
(7) For each kw = (k,w) in C2 do {
(8) kw.support ← kw.support. count (kw, p)
(9) }
(10) }
(11) $F_2 \leftarrow \left\{ kw \in C_2 \middle| kw.support \right\rangle \sigma \}$
(12) ← 2 P Filter Table(P_1 ,F_2)
(13) For (i←3; F i−1 ≠ ϕ ; i ←i +1) do {
(14) C_i ← (F_{i-1} F2)
(15) $C_i \leftarrow C_i - \left\{ kw \middle| (i-1)\text{Item} - \text{setofkw} \notin F_{i-1} \right\}$
(16) ← P_i Filter Table (P_{i-1} ,F_{i-1})
(17) For each Patches p in i P do {
(18) for each kw in C_i do
(19) {
(20) kw.support ← kw.support + count (kw, p)
(21) }
(22) }
(23) $F_i \leftarrow \left\{ kw \in C_i \middle| kw.support > \sigma \right\}$
(24) }
(25) }
(24) $sets \leftarrow \bigcup_i \{ kw \in F_i \middle| i > 1 \}$
(25) Rule=ϕ
(26) for each item set I in sets do {
(27) Rule Rule {f kw| f kw I f Is an item set kw C } 0
← + ⇒ ∪ ∈ ∧ ∧ ∈
}

The association rules are constrained such that the antecedent of the rules is composed of conjunction of features from the brain image while the consequent of the rule is always the class label to which the brain image belongs [33].

3.3.1 Pruning Techniques

The rules generated in the mining phase are expected to be very large. This could be a problem in applications where fast responses are required. Hence, the pruning techniques become necessary to eliminate the specific rules and which are conflicting with the same characteristics pointing different categories [20, 30].

This can be achieved using the following conditions:
Condition 1: Given two rules R1⇒ C and R2⇒ C , the first rule is a general rule if R1 ⊆ R2. To attain this, ordering the association rules must be done as per condition 2.

Condition 2: Given two rules R1 and R2, R1 is higher ranked than R2 if:
(1) R1 has higher confidence than R2,
(2) If the confidences are equal, support of R1 must exceed support of R2
(3) If both confidences and supports are equal, but R1 has less attributes in left hand side than R2.
Condition 3: The rules R1⇒ C1 andR1⇒ C2 , represents are conflict in nature. Based on the above conditions, duplicates have been eliminated. The set of rules that are selected after pruning represents the actual classifier. These conditions have been used to predict to which class the new test image belongs.

3.3.2 Classification

After the completion of training phase, an actual classifier with pruned set of association rules can be built for training the brain images [18] depicted in figure 2.. Each training image is associated with a set of keywords. Keywords are representative

words chosen by a specialist to use in the diagnosis of a medical image. The knowledge of specialists should also be considered during the processing of mining medical images in order to validate the results. The extracted features of the test image and the feature vector generated can be submitted to the classifier, which uses the association rules and generates set of keywords to compose the diagnosis of a test image.

Algorithm:
Input: Feature vector F of the test image,
threshold
Output: set of keywords S
Method:
(1) for each rule r∈R of the form body → head do
(2) {
(3) for each itemset h ∈ head do
(4) {
(5) if body matches F then
(6) increase the number of matches by 1
(7) Else
(8) increase the number of non matches by 1
(9) }
(10) }
// to generate keywords
(11) for each rule r∈R of the form body → head do
(12) {
(13) for each item set h ∈ R head do
(14) {
(15) if (n(Mh)/ n(Mh) + n(Nh))) ≥ T then
(16) if h∉S then
(17) add h in S
(18) }
(19) }
(20) return S

This classifier returns the multiple classes when processing a test image. The algorithm developed has been employed to generate suggestions for diagnosis. This algorithm stores all item sets (i.e. Set of keywords) belonging to the head of the rules in a data structure. An item set h is returned in the suggested diagnosis if the condition is satisfied as the given equation

$$n(Mh)/(n(Mh)+n(Nh)) \geq T \quad (15)$$

where, n(Mh) is the number of matches of the item set h and n(Nh) is the number of non-matches. Threshold T is employed to limit the minimal number of matches required to return an item set in the suggested diagnosis. A match occurs when the image features satisfy the body part of the rule.

3.3.3 System description

Overview of the proposed system is shown in Fig 2. The proposed system is mainly divided into two phases: the training phase and the test phase. Data cleaning and feature extraction are common for both the training set of mammogram images and the test set [31,32]. In the training phase, features are extracted from the images, represented in the form of feature vectors. Next, the features are discretized into intervals and the processed feature vector is merged with the keywords related with the training images [33]. This transaction representation is submitted to the MARI (Mining Association Rule in Image database) algorithm for association rule mining, which finally produces a pruned set of rules representing the actual classifier [34, 35]. In the test phase, the feature vector obtained from the test images are submitted to the classifier which makes use of the association rules to generate keywords to compose the diagnosis of the test image. These keywords have been used to classify the three categories of CT scan brain images as normal image, benign (tumor without cancerous tissues) image and malignant (tumor with cancerous tissues)
image.

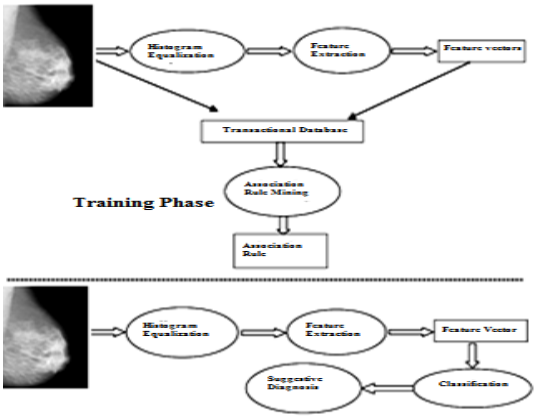

Figure2. Proposed Method for classification

3.3.4 *Performance evaluation criteria*

The confusion matrix can be used to determine the performance of the proposed method and is shown in Fig 2. This matrix describes all possible outcomes of a prediction results in table structure. The possible outcomes of a two class prediction be represented as True positive (TP), True negative (TN), False Positive (FP) and False Negative (FN). The normal and abnormal images are correctly classified as True Positive and True Negative respectively. A False Positive is when the outcome is incorrectly classified as positive (yes) when it is a negative (no). False Positive is the False alarm in the classification process. A false negative is when the outcome is incorrectly predicted as negative when it should have been in fact positive.

From the confusion matrix, the precision and recall values can be measured using the formula. Precision: It is defined as the fraction of the classified image, which is relevant to the predictions. It is represented as

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}}$$

Recall: It is defined as the fraction of the classified image for all the relevant predictions. It is given as

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}}$$

4. Statistical and texture features

Statistical and texture features are extracted for each ROI. The extracted features are then used in neural network classifier to train it for the recognition of a particular ROI of similar nature. These features are pixel value, mean, standard deviation, smoothness, entropy, skewness, kurtosis and uniformity. These are adapted from [10, 11, 19, 20]. Usually these features are extracted for each centered pixel centered on the window size 1024 ×1024. In this application, same features are extracted in a modified way. In stead of extracting each pixel, the features are extracted for the whole image by treating the whole image as a pixel.

4.1. Mean Value

The mean, μ of the pixel values in the defined window, estimates the value in the image in which central clustering occurs.

$$\mu = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)$$

Where *p(i,j)* is the pixel value at point *(i,j)* of an image of size MxN.

4.2. Standard Deviation

The standard deviation, σ is the estimate of the mean square deviation of grey pixel value *p(i,j)* from its mean value μ . Standard deviation describes the dispersion within a local region. The standard deviation is defined as:

$$\sigma = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(p(i,j)-\mu)^2}$$

4.3. Smoothness

Relative smoothness, *R* is a measure of grey level contrast that can be used to establish descriptors of relative smoothness.

$$R = 1 - \frac{1}{1+\sigma^2}$$

Where, is the standard deviation of the image.

4.4. Entropy

Entropy, *h* can also be used to describe the distribution variation in a region. Overall entropy of the image can be calculated as:

$$h = -\sum_{k=0}^{L-1} \mathrm{Pr}_k \left(\log_2 \mathrm{Pr}_k\right)$$

Where, Pr_k is the probability of the k-th grey level, which can be calculated as $Z_k / M \times N$, Z_k is the total number of pixels with the k-th grey level and L is the total number of grey levels.

4.5. Skewness

Skewness, *S* characterizes the degree of asymmetry of a pixel distribution in the specified window around its mean. Skewness is a pure number that characterizes only the shape of the distribution.

$$S = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[\frac{p(i,j)-\mu}{\sigma}\right]^3$$

Where, p(i, j) is the pixel value at point *(i,j),* , μ and , are the mean and standard deviation respectively.

4.6. Kurtosis

Kurtosis, *K* measures the peakness or flatness of a distribution relative to a normal distribution. The conventional definition of kurtosis is:

$$K = \left\{ \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[\frac{p(i,j) - \mu}{\sigma} \right]^4 \right\} - 3$$

Where, p(i, j) is the pixel value at point *(i,j),* μ and , are the mean and standard deviation respectively. The -3 term makes the value zero for a normal distribution.

4.7. Uniformity

Uniformity, *U* is a texture measure based on histogram and is defined as:

$$U = \sum_{k=0}^{L-1} Pr_k^{\,2}$$

Where, k Pr is the probability of the k-th grey level. Because the Pr_k have values in the range [0,1] and their sum equals 1, U is maximum in which all grey levels are equal, and decreases from there. Before computing any of the descriptive texture features above, the pixel values of the image were normalized by dividing each pixel by 255 in order to achieve computational consistency.

5. Proposed methods

ARM exploits the major mammographic characteristics texture to classify the mass into benign or malignant. To fulfill the objectives of this paper ARM uses mean, standard deviation, entropy, skewness, kurtosis and uniformity features that is described in section 4. These 7 features are used in preparing the training data which are obtained from the whole extracted mass region. To do that an image file of mass (benign or malignant) is loaded and the 'Features Calc" button is pressed that is shown in figure 4. Seven features of the loaded image is calculated. Based on the prior information either benign (0) or malignant (1) button is selected and corresponding 0 or 1 is placed in the output field. The calculated 7 features and their corresponding target value (for benign=0 and malignant=1) are stored in a file. The same process is repeated for more masses both malignant and benign. This way all the training samples are stored in the file that is used as inputs to the ARM and based on the support and confidence the rule is generated for testing the classifier.

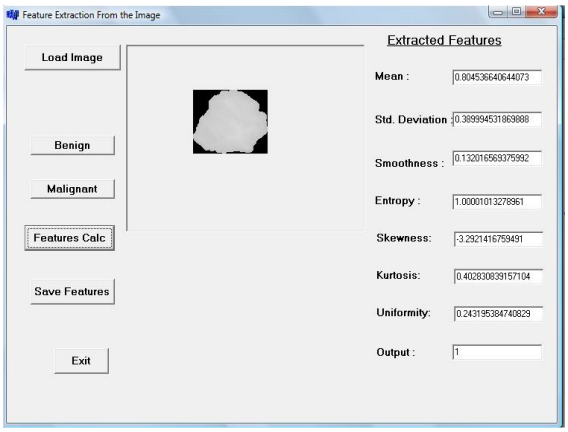

Figure 4: Training data preparation

Once the training samples are available, the rules are ranked by score, generality, and support. A decision list is prepared taking the rank value and secondly the weights are combined for class value prediction. These samples are used to train the classifier and the generated rule is applied to the testing data for classification.

6. Simulation results and performance evaluation

6.1. Image Database

To develop and evaluate the proposed system we used the Mammographic Image Analysis Society (Mini MIAS) [21] database. It is an organization of UK research group. Films were taken from UK National Breast Screening Program that includes radiologist's "truth"- markings on the locations of any abnormalities that may be present. Images are available online at the Pilot European Image Processing Archive (PEIPA) at the University of Essex. This database contains left and right breast images for a total of 161 (322 images) patients with ages between 50 and 65. All images are digitized at a resolution of 1024x1024 pixels and at 8-bit grey scale level. The existing data in the

collection consists of the location of the abnormality (like the center of a circle surrounding the tumor), its radius, breast position (left or right), type of breast tissues (fatty, fatty-glandular and dense) and tumor type if it exists (benign or malign).

Each of the abnormalities has been diagnosed and confirmed by a biopsy to indicate its severity: benign or malignant. In this database, 52 images contain abnormalities (malignant masses) and 106 images are classed as normal and rest of them either contains microcalcifications or benign.

6.2. Results and Performance

In this paper we used association rule mining [22-28] for the classification of mammograms. The average accuracy is 93.91 %. We have used the precision and recall measures as the evaluation metric for mammogram classification. Precision is the fraction of the number of true positive predictions divided by the total number of true positives in the set. Recall is the total number of predictions divided by the total number of true positives in the set. The testing result using the selected features is given in table 1. The selected features are used for classification. For classification of samples, we have employed the freely available Machine Learning package, WEKA [29]. Out of 115 images in the dataset, 76 were used for training and the remaining 39 for testing purposes. The overall accuracy is approximately 94%.

Table 1: Results obtained by proposed method (correctly classified)

Malignant	95.23%
Benign	92.30%

The confusion matrix has been obtained from the testing part .In this case for example out of 97 actual malignant images 09 images was classified as normal. In case of benign and normal all images are correctly classified. The confusion matrix is given in Table 2.

Table 2: Confusion matrix

	Predicted class	
Actual	Benign	Malignant
Benign	60	3
Malignant	04	48

The classification performance can be assessed in terms of the sensitivity and specificity of the system. Sensitivity (SN) is the proportion of actual positives which are correctly identified and it is mathematically defined in equation 1 and specificity (SP) is the proportion of negatives which are correctly identified and is mathematically defined in equation 2. The sensitivity and specificity of the proposed system compared to the radiologist's sensitivity is shown in table 3.

$$SN = \frac{TP}{TP + FN} \quad (1)$$

$$SP = \frac{TN}{TN + FP} \quad (2)$$

Table 3: Confusion matrix

TP	TN	FP	FN	SN (%)	SP (%)	Radiologist Sensitivity (%)
48	60	03	04	92.30	95.23	80

7. Conclusions and Future work

Mass classification is a vital stage for the performance of the computer aided breast cancer detection. It reduces the false positive rate by reducing the unnecessary biopsy and health care cost as well. Different classifiers were used in breast cancer detection from mammograms. However, ARM shows very good performance in diagnostic systems. In this paper, a correlated association rule mining is proposed for mass classification. The performance of the proposed structure is evaluated in terms of efficiency, adaptability and robustness. Computational time is around 15~20 ms for each mass classification. It was evaluated on 115 images containing malignant and benign masses with different size, shape and contrast. The algorithm works properly in all cases and the proposed structure was evaluated with and without preliminary de-noising steps. In both cases results are found to be

comparable. Using the proposed ARM classifier, 92. 30% sensitivity and 95.23% specificity is achieved which is very much encouraging compare to the radiologist's sensitivity 80%. Though emphasis has been given on classification rule but feature selection is also very vital for obtaining accuracy because of under fitting either or over fitting to the classifier. Subsequently feature optimization and hybrid feature selection technique will probably best suit to classifier for better accuracy. In addition to that hybrid classifier models may be employed for classification with appropriate methodology so that time complexity as well as execution for detection of result will be well within the acceptable range. Association rule can be further developed to correlate the feature values accurately with high confidence and support. Gray level coocuurence matrix (GLCM) features are also very important for classification accuracy and that to be more précised by optimization methods. We shall workout on those mentioned features as well as most significant features to be well correlated and better optimized and suitable for classifier to classify with minimum time complexity and higher accuracy.

References

[1] A.Oliver, J. Freixenet, R. Marti, J. Pont, E. Perez, E.R.E. Denton, R. Zwiggelaar, (2008) "A novel breast tissue density classification methodology", IEEE Transactions on Information Technology in Biomedicine, Vol. 12, No.1, pp. 55-65.

[2] R.E. Bird (1990) "Professional quality assurance for mammography screening programs", Journal of Radiology, Vol. 175, pp. 587-605.

[3] H.D. Cheng, X.J. Shi, R. Min, L.M. Hu, X.P. Cai, H.N. Du (2006) "Approaches for automated detection and classification of masses in mammograms", Pattern Recognition, Vol. 39, pp. 646-668.

[4] S.C. Yang, C.M. Wany et.al. (2005) "A Computer-aided system for mass detection and classification indigitized mammograms", Journal of Biomedical Engineering- Applications, Basis and Communications, Vol. 17, pp. 215-228.

[5] R. O. Duda, P. E. Hart, D. G. Stork (2001), Pattern Classification, John Wiley and Sons, second edition.

[6] H.P. Chan, D. Wei, M.A. Helvie, B. Sahiner et.al. (1995) "Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space", Journal of Physics in Medicine and Biology, Vol. 40, pp. 857-876.

[7] B. Sahiner, N. Petrick, H.P. Chan (2001) "Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization", IEEE Trans. Med Imaging, Vol. 20, No. 12, pp. 1275–1284.

[8] J.L. Viton, M. Rasigni, G. Rasigni, A. Liebaria (1996) "Method for characterizing masses in digital mammograms", Opt. Eng., Vol. 35, No. 12, pp. 3453–3459.

[9] P.J.G. Lisboa, (2000) "A review of evidence of health benefits from artificial neural networks in medical intervention", Neural Networks, Vol. 15, pp. 11-39.

[10] Y. Alginahi (2004) "Computer analysis of composite documents with non-uniform background", PhD Thesis, Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada.

[11] Y. Alginahi (2008) "Thresholding and character recognition in security documents with watermarked background", Proc. Int. Conf. on Digital Image Computing: Techniques and Applications, pp. 220-225.

[12] B. Zheng, Y.H. Chang, X.H. Wang, W.F. Good (1999) "Comparison of artificial neural network and Bayesian belief network in a computer assisted diagnosis scheme for mammography", IEEEInternational Conference on Neural Networks, pp. 4181–4185.

[13] B. Sahiner, H.P. Chan, N. Petrick, M.A. Helvie, M.M. Goodsitt (1998) "Design of a high-sensitivity classifier based on a genetic algorithm: application to computer-aided diagnosis", Phys. Med. BiolVol. 43, No. 10, pp. 2853–2871.

[14]. Agrawal R, Imielinski T, Swami AN. Mining association rules between sets of items in large

databases. Proceedings of the 1993 ACM SIGMOD ICMD, ACM; Washington, D.C.. 1993. pp. 207–216
[15]*1.* Ribeiro MX, Traina AJM, Balan AGR, Traina C, Jr, Marques PMA. SuGAR: A framework to support mammogram diagnosis. IEEE CBMS 2007; Maribor, Slovenia. 2007. pp. 47–52.
[16]. Tseng SV, Wang M-H, Su J-H. A new method for image classification by using multilevel association rules. Presented at ICDE 05; Tokyo. 2005. pp. 1180–1187.
[17]. Yun J, Zhanhuai L, Yong W, Longbo Z. Joining associative classifier for medical images. *HIS* 2005
[18] S. T. Vincent, W. Ming-Hsiang, and S.J. Hwung, "A New Method for Image Classification by Using Multilevel Association Rules," In Proc: 21st International Conference on Data Engineering Workshops (ICDEW), 2005, pp.1180-1188.
[19] Y. Alginahi, M.A. Sid-Ahmed and M. Ahmadi (2004) "Local thresholding of composite documentsusing Multi-layer Perceptron Neural Network", in 47th IEEE International Midwest Symposium on Circuits and Systems, pp. 209-212.
[20] B. Liu, W. Hsu, Y. Ma, " Pruning and Summarizing the Discovered Associations," In Proc: ACM SIGKDD International Conference on Knowledge Discovery & Data Mining , 1999, pp. 81-105.
[21] J. Suckling et al. (1994) "The Mammographic Image Analysis Society Digital Mammogram Database Excerpta Medica", International Congress Series, Vol. 1069, pp. 375-378.
[22] X Wang, M R Smith, R M Rangayyan "Mammographic information analysis through association-rule mining,Canadian Conference on Electrical and Computer Engineering 2004 (2004), DOI: 10.1109/CCECE.2004.1349689
[23]Roselin, R., Thangavel, K "Classification ensemble for mammograms using Ant-Miner" International Conference on Computing Communication and Networking Technologies (ICCCNT), 2010 DOI:10.1109/ICCCNT.2010.5592607 pp.1 – 6
[24]Sumeet Dua,. Harpreet Singh, H.W. Thompson," Associative Classification of Mammograms using Weighted Rules" Expert Syst Appl. 2009 July 1; 36(5): 9250–9259doi:10.1016/j.eswa.2008.12.050
[25]P.Rajendran,M.Madheswaran, "Novel Fuzzy Association Rule Image Mining Algorithm for Medical Decision Support System ",International Journal of Computer Applications 1(20): 87-94,2010, DOI:10.5120/415-613
[26]Lukasz Kobyli´nski, Krzysztof Walczak," Image Classification with Customized Associative Classifiers", Proceedings of the International Multiconference on Computer Science and Information Technology pp. 85–91
[27]Benaki Lairenjam, Siri Krishan Wasan. Neural Network with Classification Based on Multiple Association Rule for Classifying Mammographic Data. In Proceedings of IDEAL'2009. pp.465~476
[28]K. Thangavel, A. Kaja Mohideen," Classification of Microcalcifications Using Multi-Dimensional Genetic Association Rule Miner", International Journal of Recent Trends in Engineering, Vol 2, No. 2, November 2009
[29]. Holmes, G., Donkin, A., Witten, I.H.: WEKA: a machine learning workbench. In: Proceedings Second Australia and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia, pp. 357-361, 1994
[30] R.Z. Osmar, L.A. Maria," On Pruning and Tuning Rules for Associative Classifiers," In Proc: 9th International Conference(KES) ,K-Based-Intelligent Information and Engineering Systems Melbourne, part III, 2005.
[31] B.A. Dogu, H. Markus, A.Tuukka, D. Prasun, and H. Jari , "Texture Based Classification and Segmentation of Tissues Using DTCWT Feature Extraction Methods," In Proc: 21st IEEE International Symposium on Computer-Based Medical Systems, 2008, pp.614-619.
[32] P. Dollar, T. Zhuowen, T. Hai, and S. Belongie, " Feature Mining for Image Classification," In Proc: IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1-6.
[33] J. Dougherty, R. Kohavi, and M. sahami, "Supervised and Unsupervised Discretization of continuous features," In Proc: 12th International Conference Machine Learning, 1995, pp.56-69.
[34] R. Agrawal, T. Imielinski, And A.N. Swami, "Mining association rules between sets of items in large databases," In Proc: ACMSIGMOD Int. Conf. Manage, Washington, DC, 1993, pp. 207-216.

[35] S. Kotsiantis, D.Kanellopoulos, "Association Rules Mining: A Recent Overview," GESTS International Transactions on Computer Science and Engineering, 32 (1):2006, pp. 71-82.
[36] Joshi, K., Kumar, V., Sundaresan, V., Ashish Kumar Karanam, S., Dhabliya, D., Daniel Shadrach, F., & Ramachandra, A. C. (2022). Intelligent fusion approach for MRI and CT imaging using CNN with wavelet transform approach. Paper presented at the IEEE International Conference on Knowledge Engineering and Communication Systems, ICKES 2022, doi:10.1109/ICKECS56523.2022.10060322 Retrieved from www.scopus.com
[37] Juneja, V., Singh, S., Jain, V., Pandey, K. K., Dhabliya, D., Gupta, A., & Pandey, D. (2023). Optimization-based data science for an IoT service applicable in smart cities. Handbook of research on data-driven mathematical modeling in smart cities (pp. 300-321) doi:10.4018/978-1-6684-6408-3.ch016 Retrieved from www.scopus.com