Research Article
Effective Rule Based Classifier using Multivariate Filter and Genetic Miner for Mammographic Image Classification

Nirase Fathima Abubacker, Azreen Azman, Masrah Azrifah Azmi Murad and Shyamala Doraisamy
Faculty of Computer Science and Information Technologi, Universiti Putra Malaysia, Serdang, Malaysia

Abstract: Mammography is an important examination in the early detection of breast abnormalities. Automatic classifications of mammogram images into normal, benign or malignant would help the radiologists in diagnosis of breast cancer cases. This study investigates the effectiveness of using rule-based classifiers with multivariate filter and genetic miner to classify mammogram images. The method discovers association rules with the classes as the consequence and classifies the images based on the Highest Average Confidence of the association rules (HAvC) matched for the classes. In the association rules mining stage, Correlation based Feature Selection (CFS) plays an enormous significance to reduce the complexity of image mining process is used in this study as a feature selection method and a modified genetic association rule mining technique, the GARM, is used to discover the rules. The method is evaluated on mammogram image dataset with 240 images taken from DDSM. The performance of the method is compared against other classifiers such as SMO; Naïve Bayes and J48. The performance of the proposed method is promising with 88% accuracy and outperforms other classifiers in the context of mammogram image classification.

Keywords: Association rule mining, correlation-based feature selection, mammographic image classification, multivariate filters

INTRODUCTION

In Western and developing countries, including Malaysia, breast cancer has become one of the most important health problems. It is possible to detect breast cancer early when regular screening examinations are done. Mammography is the most excellent way to detect the breast cancer in its treatable and early stage. It is a special type of x-ray photograph that uses high-resolution film, high contrast and low dose x-ray for imaging the breasts. Unfortunately, not all breast cancers can be detected by mammogram. In addition, the diagnosis of all types of breast diseases depends on a biopsy and the decision to conduct biopsy procedure depends on the mammographic findings. However, 10-30% of breast cancer cases are missed by mammography, which leads to a delay in diagnosis and could result in more radical surgery or could be fatal. Alternatively, Computer Aided Diagnosis (CAD) technology can assist radiologists in their interpretation of mammogram images to reduce the number of missing cases. In this case, classification techniques can be used to automatically classify the mammogram images into normal, benign or malignant. As such, it will help the radiologists in their diagnosis of breast cancer.

In this context, linear discriminate analyses used by Pang *et al.* (2005), are considered to be the traditional classification method, but have shown poor performance for nonlinear separable data as compared to linear separable data. A few other researchers have used different classifiers such as using rough sets by Abu-Amara and Abdel-Qader (2009), Bayesian Artificial Neural Network (ANN) by Lisboa (2002) for the classification of masses. Among these classifiers, Association Rules mining (AR) gains popularity as the classification technique for this problem due to its nature in reflecting close dependencies among features in composing rules. However, one of the main challenges in using AR is the generation of a huge number of rules based on the frequent item set generated for all features. The number of rules grows exponentially with the number of items. As such, there will be many rules generated for each class. Similar to other data mining approaches, the performance of AR depends largely on the features used for the mining task. A huge number of features used may enhance the classification accuracy, but will affect the efficiency of AR. Hence, it is important to reduce the number of features through the process of feature selection and discovering high-level prediction rules to improve the performance of the classification.

This study aims to build an effective associative classifier with the use of Correlation-based Feature Selection (CFS) as the feature selection method,

together with the Genetic Association Rule Mining (GARM) that allows interactions among features and performs global search in discovering the rules, by using a fitness function as threshold value to evaluate the rules for mammography image classification.

LITERATURE REVIEW

Data Mining is a powerful tool in revealing hidden relationships of a large database. A descriptive data mining task called Association Rule Mining summarizes the database information and analyses data relationships for extracting patterns. The quality of decision for automatic diagnosis in the medical field depends on the quality data. The elimination of redundant features reduces the data size. Integrating feature selection reduces number of features, removes noisy or irrelevant data, thus speeding up the mining process. Hence, mining on a reduced set of data helps to make the association rule pattern to be discovered easily and to improve its predictive accuracy. There are two types of data reduction methods, which are wrapper and filter methods. Even though wrapper methods can produce better result, they are expensive for the large dimensional database. On the other hand filter method is computationally simple and fast and precedes the actual association rule generation process. Filter methods use some properties of the data to select the feature. An intrinsic property such as entropy has been used as a filter method for feature selection. Yu and Liu (2003) proposed a selection method using a correlation measure that identifies redundant features. Many popular search procedures like particle swarm optimization, sequential forward selection, sequential backward selection, genetic search, etc. have been proposed in many researches. Genetic Algorithms are effectively applied to a diversity of problems like feature selection problems by Barlak (2007), data mining problems by Ishibuchi and Yamamoto (2004), scheduling problems by Gonçalves *et al*. (2005), multiple objective problems by Deb *et al*. (2000) and Dias and De Vasconcelos (2002), traveling salesman problems by Tsai *et al*. (2002). Since the univariate filters does not justify for interactions between features multivariate filter Correlation based Feature Selection (CFS) can be used to overcome the drawback of the univariate filter for determining the best feature subset.

A few authors have proposed many algorithms in recent years for mining association rules such as Genetic Algorithm (GA). Saggar *et al*. (2004) have used the GA to optimize the rules generated by the Apriori algorithm. The rule based classification systems can even predict negative rules with the improvements applied to GAs. Ghosh and Nath (2004) solved the multi-objective rule mining problems by representing the rules as chromosomes using Michigan approach. Shrivastava *et al*. (2010) extracted interesting association rules using an optimized GA using the measures like interestingness, completeness, support and confidence. Jain *et al*. (2012) used a genetic algorithm for the whole process of optimization of the rule set to find a reduction of Negative and Positive Association Rule Mining. Lim *et al*. (2012) proposed Hybrid genetic algorithm for mining workflow best practices, using correlation measures instead of traditional support and confidence. Wakabi-Waiswa *et al*. (2011) used a genetic algorithm for a structured method to find the unknown facts in large data sets. Nahar *et al*. (2013) proposed Association rule mining for the detection of sick and healthy heart disease factors, using the UCI Cleveland dataset, a biological database. Pang *et al*. (2005) proposed a method of utilizing linkage among feature selections using Multi-objective Genetic Algorithm for data quality mining. Lee *et al*. (2013) used Genetic algorithm for association rule generation related to hypertension and diabetes in discovering medical knowledge for young adults with acute myocardial infarction. Keshavamurthy *et al*. (2013) compared conventional mining algorithm, i.e., Apriori algorithm with the proposed genetic algorithm in local search for privacy preserving over distributed databases. In the Apriori algorithm population is formed in only single recursion, but in genetic algorithm population is formed in every new production. In order to overcome the disadvantages of Apriori algorithm, Genetic algorithm can be used for association rule mining. Above research gap helps to improve the mining process using Genetic algorithm, which can further improve the classification accuracy.

Leonardo *et al*. (2009) presented a methodology or detection of masses on DDSM mammograms. They used segmented images using the K-means algorithm and described the texture of the segmented structures using co-occurrence matrix to accomplish classification through Support Vector Machines that gained 85% of accuracy. Bovis and Singh (2002) investigated a method to classify mammograms based on the prior knowledge about breast type. By utilizing this knowledge their application aims to increase the sensitivity of detecting breast cancer. They used feed-forward Artificial Neural Network (ANN) that comprises multiple classifiers for the classification purpose. The data set used in this study is the Digital Database of Screening Mammograms (DDSM) that gives an average recognition rate on a test of 71.4%. Cascio *et al*. (2006) presented an approach for detecting mammographic lesions. They used ROI Hunter algorithm for surface reduction with no loss of meaningful information and used the output neuron probability obtained from supervised neural network to classify the ROI pathology. Eddaoudi *et al*. (2011) proposed SVM classification and texture analysis for mass detection. The results obtained with original mammograms, showed a classification rate of 77% in

average. Singh (2011) proposed a method that uses the query image with stored images in database for tissue classification and analysis of breast images. The accurate result was obtained to be 85.7% and showed that the suggested features can be used for both classification and retrieval of mammography images. Mavroforakis *et al.* (2006) has achieved an optimal classification of mammograms with a score of 83.9%, through SVM classifiers for the breast mass. Abu-Amara and Abdel-Qader (2009) proposed a hybrid mammogram classification using rough set and fuzzy classifier. They attempt to reduce the effect of data inconsistency using Rough set model and used a fuzzy classifier for labeling normal or abnormal regions that produced an accuracy of 84.03%. Among all these, the classifiers that involve association rules induced from significant event associations are more easily understood by humans as it brings close dependencies among features for composing rules. Another advantage of this approach is its greater flexibility in handling unstructured data as it provides confidence probability for solving classification problem of uncertainty. Hence associative classifiers that use data mining techniques have become a hot theme in recent years.

METHODOLOGY

The framework consists of two phases, training phase and testing phase. The feature vector is formed Euclidean distance for the comparison of the features of using the features that are extracted from the training images and the diagnosis report that are extracted from ground truth provided by Heath *et al.* (2000) for mammogram database. The genetic association rule miner takes this input to generate association rules and will be stored in a database. For a new test image, the same features are extracted and the feature vector is given as an input to the classifier HAvCBC that uses the high average confidence association rules based on category obtained from training images for an effective classification of mammograms. The framework of the proposed method is given in Fig. 1.

Step 1: Feature extraction: The features that are used to classify the normal and abnormal lesions can be represented as mathematical descriptions. In general, several features are usually used to express the characteristics of an image. Feature extraction methodology analyses mammogram images to extract the most prominent features that represent various classes of the images. Unlike the complicated process of an individual person like radiologist to classify a mass, the decision on classifications can be made by machines with few limited features. In this study, the statistical texture features such as contrast, coefficient, entropy, energy, homogeneity and a few otherfeatures proposed by Haralick *et al.* (1973) and Soh and Tsatsoulis (1999) that efficiently classify the benign
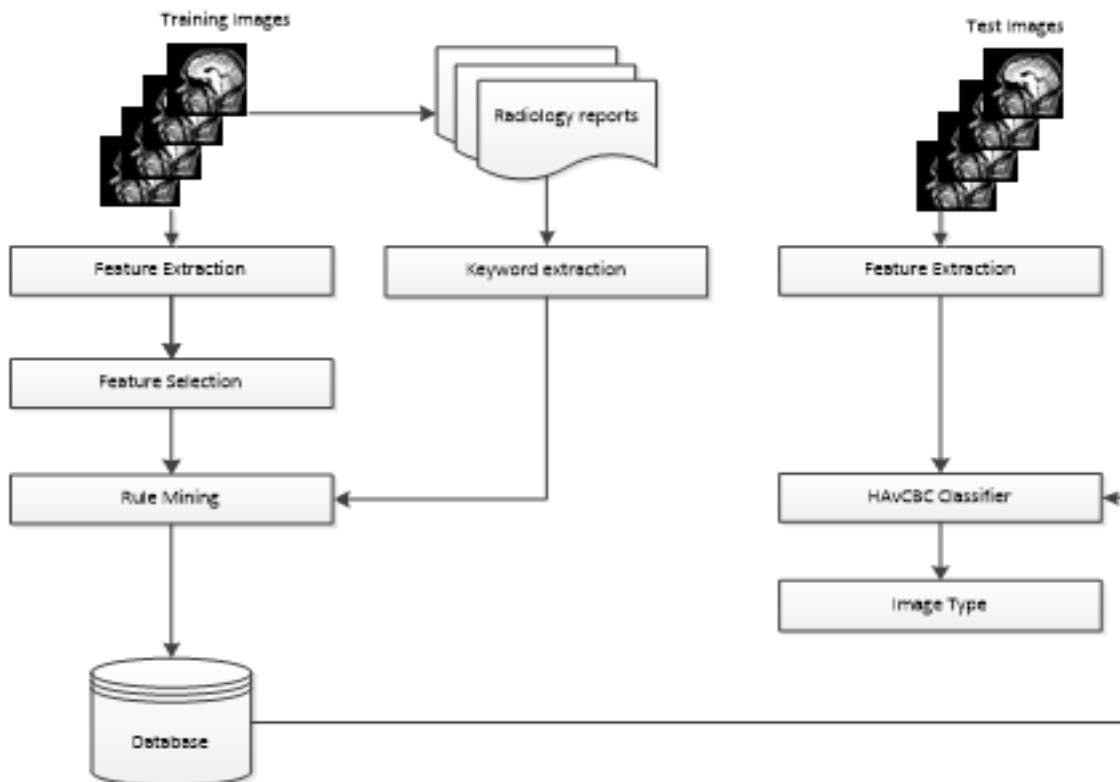

Fig. 1: Proposed system framework

and malignant mammograms are extracted with the distance between the pixel of interest and its neighbor equal to 1 and the angle of 0. Let p (i, j) be the (i, j)th entry in a normalized Gray Level Co-occurrence Matrix (GLCM). μ and are the mean and standard deviation for the rows and columns of the matrix. Twenty features that were derived from GLCM are Auto Correlation, contrast, energy, homogeneity, correlation, inverse difference moment normalized, inverse difference normalized, entropy, difference entropy, maximum probability, shade, dissimilarity, variance, prominence, information correlation 1, information correlation 2, sum of squares, sum average, sum entropy.

Step 2: Correlation-based Feature Selection (CFS):
Many of the features extracted during the training phase are either partially or completely irrelevant to an object that has no effect on the target concept. The representation and quality of data may have an effect on a given task in machine learning. Liu and Motoda (2007) the feature selection process helps to remove redundant and irrelevant features, thus reducing the feature space. This minimizes the computation time and helps in improving the prediction accuracy. Data reduction methods are of two types, wrapper and filter methods. On the other hand filter approach measures the feature subset relevance and is independent of learning algorithm. Since the univariate filters does not justify for interactions between features we use multivariate filter Correlation based Feature Selection (CFS) to determine the best feature subset. CFS involves heuristic search to evaluate the subset of features. This is a simple filter algorithm based on correlation heuristic function. This function evaluates the subsets that are highly correlated and uncorrelated with each other and class. The features are accepted depending on its level of extent in predicting its class. The features that did not influence the class will be ignored as irrelevant features.

The features extracted for the digital mammograms are often contain many continuous attributes. The method of transforming the attributes with continuous value into nominal value is called Discretization. This study uses an unsupervised discretization method proposed by Hall (1999) and Dougherty *et al.* (1995) that places values in the same interval that are strongly associated. The values of all features are sorted at the first level and divided into n intervals, then for each interval of a class that has the strong majority an interval boundary is created. The discretized values are stored in a database, in which, each attribute (feature) is represented by one column with the class attribute that is represented by the last column and tuples are used to represent images.

The discretized data are then passed to CFS. For the prediction of class labels, this CFS considers the efficacy of individual features and its inter-correlation. If given the inter-correlation between each pair of features and the correlation between each of the features and its class, then the correlation between a subset of features selected for evaluation can be predicted from the formula:

$$cr_{zc} = \frac{f \overline{cr_{zi}}}{\sqrt{f + f(f-1)\overline{\overline{cr_{ii}}}}}$$

where, cr_{zc} is heuristic merit of a feature subset for f number of features, $c \; \overline{r_{zi}}$ is the average of the correlations between the class and the features and $c \, \overline{r_{ii}}$ is average inter-correlation between feature pairs. The subset with the highest cr_{zc} value is used to reduce the data dimensionality.

The feature selection process together with CFS undergoes some search procedure. This study uses GA as a search method with CFS as a subset evaluating mechanism. Genetic Algorithm (GAs) is modeled based on the process of natural selection. At every iteration new populations are generated from old ones in each iteration. They are actually binary encoded strings. Every string is evaluated to measure its fitness value for the problem. Likewise the entire generation of new strings can be computed using the genetic operators, on an initially random population. This operative way of discovering large search space is essential for feature selection. The individual fitness can be decided by the correlation between the features. Based on the correlation coefficient the individuals will be assigned a rank by the fitness function. The features that have the lower correlation coefficients and with higher fitness value will be appropriate for crossover operations. The features that were selected in this process are Auto_correlation, Entropy Contrast, Cluster_Shade, Sum_Entropy, Sum_of_Squares, Sum_Variance, Inf_Measure_Corr1, Inf_Measure_Corr2.

Step 3: Rule mining: For each image, the selected discretized features are stored in such a way that n columns represent n features while the last column represent a class (e.g., normal, benign and malignant).A modified genetic association rule mining, GARM is used to discover the frequent item set and eventually to generate the rules. Genetic Algorithm (GA), can generate good rules by performing a global search with better attribute interactions, thus can improve the effectiveness of association rule mining. For each category in the database, the GA is applied separately to construct sets of rules. For each rule the fitness value is calculated and the rule that has the highest fitness value in each population will be stored as the global best rule. Then, the best rules from each category are pooled to form rule set. New population can be computed using

the genetic operators like reproduction, crossover and mutation to extract the best local rules.

Algorithm 1:

Input: A feature set f_i of the form $\{c_i, f_1, f_2 \ldots f_n)$ where c_i is the image category.

Output: Set of association rules $f_i \dashrightarrow c_i$.

Step 1: Load the selected feature subset $f_1, f_2, f_3, f_4, f_5, \ldots\ldots f_n$

Step 2: Construct N chromosomes (rules) for M populations for each category

Step 3: Calculate the fitness value for each chromosome and rank them

Step 4: Select the higher ranked fitness value and set it as best chromosome

Step 5: Construct new population using genetic operators and select best LOCAL chromosome

Step 6: Compare global and local chromosomes, if global chromosomes are better than the local, continue with the older population and perform rule pruning else continue with new population

Step 7: Form association rules by decoding the best global chromosomes

Step 4: Classifier (HAvCBC): The extracted set of rules represents the actual classifier. It classifies a new image to its category. When a new image is provided, a feature vector is extracted and it searches in the rules for matching classes. Based on the matched rules in each class, the average confidence score is calculated. The class for the new image is identified based on the highest average confidence score in the class and the number of rules matched. The algorithm describes the classification of a new image.

Algorithm 2:

Input: Number of category C, list of training rules for each category Ci, Number of rules n in each category Cj, Total number of rules N.

Output: Category attached to the new image.

```
for each category Ci
  for each rule R do
    If R matches I then
      MC++
else
    NMC++

    end if
 end for
end for
for each category Ci subset
```

find $P = \frac{MC}{n}$ the percentage of set of rules that match I
find average confidence Q of rules
end for

Put the new image in the category Ci that has the highest average confidence and higher percentage of matching.

EXPERIMENTAL RESULTS

The proposed classifier takes the extracted and selected features as input to perform the classification task. For comparison, the performance of the proposed classifier is evaluated against other state of the art classifier namely Radial Basis Function (RBF) proposed by Park and Sandberg (1991), SMO by Platt (1998), Naive Bayes by Rish (2001), OneR by Holte (1993) and J48 by Rossquinlan (1993). RBF network is an artificial neural network where the radial basis functions are used as activation functions. SMO is a new Support Vector Machine (SVM) learning algorithm that solves quadratic programming problem that arises during the training of SVM. The Naïve Bayes classifier is based on the Bayes rule of conditional probability for class label prediction. OneR, short for "One Rule", is a simple, classification algorithm that generates one rule for each predictor in the data, then the smallest total error will be selected as its "one rule". J48 is a classifier that concept of information entropy to build labeled training data as decision trees. It uses the fact that by splitting the data of each attribute into smaller subsets the decisions can be made. By applying various classifiers the dataset is analyzed.

Dataset and selection of ROI: The data set used in this experiment is taken from the digital database for mammography from the University of South Florida (Vafaie and De Jong, 1992), which is DDSM. All images are digitized using LUMISYS Scanner at a resolution of 50 microns and at 12 bit grayscale levels. The dataset consists of 240 images that include three categories of which 80 are normal, 80 are benign and 80 are malignant. Then, the Region of Interest (ROI) is isolated within those images as the preprocessing step. We use the contour supplied together with the images in the DDSM dataset to extract ROIs of size 256×256 pixels.

A total of 240 ROIs is extracted with the mass centered in a window of size 256×256 pixels, where 162 are abnormal ROIs (circumscribed masses, speculated masses, ill-defined masses and architectural distortion) and 80 are normal ROIs (Fig. 2).

The performance of each classifier is measured as accuracy, precision, recall and F-Measure. Also by noticing the errors made by a classification model, the ability of the model for the correct prediction of the classes can also be identified. The confusion matrix
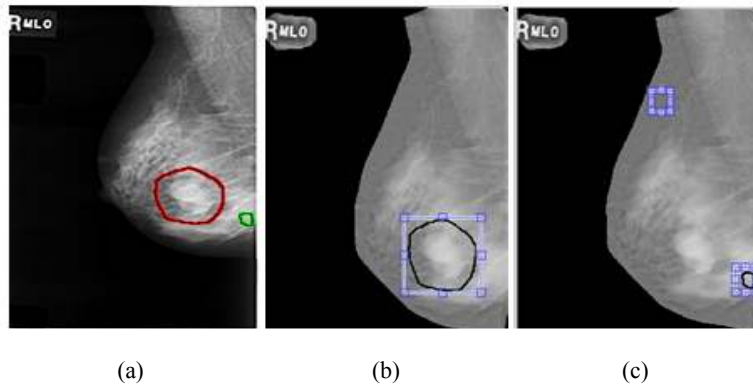
Fig. 2: Example of ROI for (a) original image, (b) abnormal, (c) normal cases

Table 1: The confusion matrix for multi class evaluation

	Predicted data		
Known data	Normal (A)	Benign (B)	Malignant (C)
Normal (A)	tp_A	e_{AB}	e_{AC}
Benign (B)	e_{BA}	tp_B	e_{BC}
Malignant (C)	e_{CA}	e_{CB}	tp_C

Table 2: The average measures for classification with different classifier

Classifier	Correct (%)	Incorrect (%)	Precision	Recall	F-measure
RBF	79.17	20.83	0.794	0.792	0.793
SMO	81.67	18.33	0.819	0.817	0.817
Naives bayes	73.33	26.67	0.730	0.733	0.731
1R	70.00	30.00	0.698	0.700	0.683
J48	77.50	22.50	0.782	0.775	0.777
HAvCBC	88.12	11.88	0.889	0.881	0.881

shown in Table 1 gives the complete representation for this multi class problem with the three classes Normal (N), Benign (B) and Malignant (M).

The matrix given in Table 1 shows the predictions for three different classes. The rows and columns of the matrix represent the known and the predicted data made by the model. The diagonal elements tp_A, tp_B, tp_C are the number of true positive (correct) classifications made for normal, benign and malignant classes. The off-diagonal elements of the matrix e_{AB}, e_{AC}, e_{BA}, e_{BC}, e_{CA}, e_{CB} show the number of incorrect (error) classifications made for each class. In the classification stage, the class for each image in the test group is identified and the average Accuracy measure (AC) for correct and incorrect classification, precision is the proportion of the predicted positive cases that were correct, recall is the is the proportion of positive cases that were correctly identified, F-Measure is the harmonic mean of precision and recall and are given by the formula (1) to (4):

$$Accuracy(AC) = \frac{tp_A + tp_B + tp_C}{N} \quad (1)$$

$$PrecisionA = \frac{tp_A}{tp_A + e_{BA} + e_{CA}} \quad (2)$$

$$RecallA = \frac{tp_A}{tp_A + e_{AB} + e_{AC}} \quad (3)$$

$$F - Measure = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

The Accuracy measurements (AC) for correct and incorrect classification for different classifiers are depicted in Table 2. The result shows that a marginal increase in accuracy for correct classification and a marginal decrease in accuracy of incorrect classification can be achieved using the proposed classifier HAvCBC. Figure 3 shows the comparison of accuracy for correct and incorrect classification with different classifiers. Figure 4 shows the comparison of precision, recall and F-Measure for different classifiers in each group. Based on the Fig. 4, it clearly shows that our proposed classifier HAvCBC reaches higher values of precision that intuitively shows the increase in the ability of the proposed classifier to correctly predict a positive a sample that is positive compared to other well-known classifiers. Precision increases for experiments with associative classifier because a system trained on association rules can cope better with the classification problem as it reveals hidden information between features and class.

Also the marginal improvement in the recall measure intuitively shows the increase in the ability of the proposed classifier to correctly identify all the positive samples. The increase in F-measure value compared to other classifiers, it enforces a better
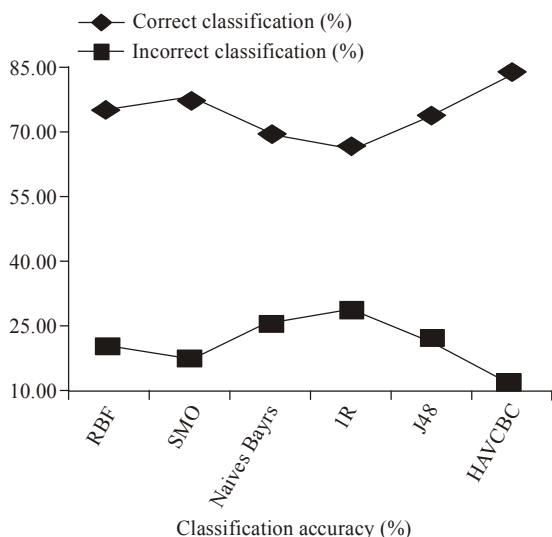
Fig. 3: Comparison of AC of correct and incorrect classification with different classifier
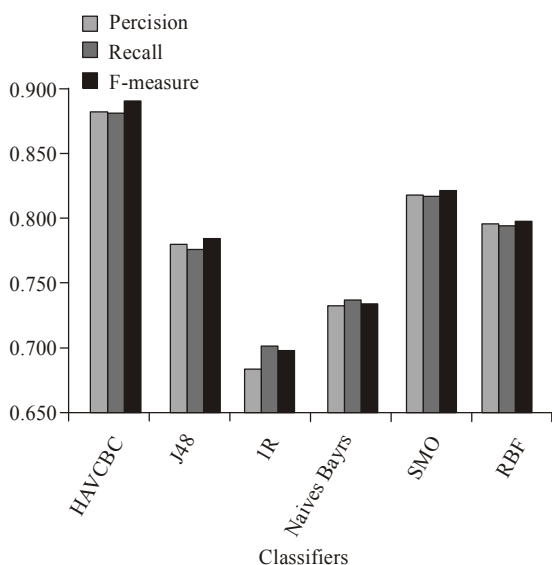

Fig. 4: Comparison of precision, recall, F-measure with different classifier

balance between performance son relevant and irrelevant labels using the proposed classifier.

CONCLUSION

In this study we proposed HavCBC, a new method that employs a correlation feature subset and genetic association rule mining for mammogram classification as an aid to CAD. The result is promising with high accuracy in its correct classification (88%) when compared to other famous classifiers (RBF, SMO, Naive Bayes, 1R, J48). Moreover the method interprets the precision value as a marginal increase in the ability of the proposed classifier not to label a sample that is negative as positive, which is required in the medical domain, to spot true positives accurately. It is learnt that the classification performance of a classifier can be more effective when the classifier is developed using the descriptive genetic association rule mining that reveals the hidden relationships between the features that helps in test phase by presenting a strong power of generalization.

REFERENCES

Abu-Amara, F. and I. Abdel-Qader, 2009. Hybrid mammogram classification using rough set and fuzzy classifier. Int. J. Biomed. Imaging, 2009: 12.

Barlak, E.S., 2007. Feature selection using genetic algorithms. Thesis, Retrieved from: http://research.sabanciuniv.edu/8518/.

Bovis, K. and S. Singh, 2002. Classification of mammographic breast density using a combined classifier paradigm. Proceeding of the Medical Image Understanding and Analysis (MIUA) Conference, Portsmouth.

Cascio, D., F. Fauci, R. Magro, G. Raso, R. Bellotti, F. De Carlo, E.L. Torres *et al.*, 2006. Mammogram segmentation by contour searching and mass lesions classification with neural network. IEEE T. Nucl. Sci., 53(5): 2827-2833.

Deb, K., S. Agrawal, A. Pratap and T. Meyarivan, 2000. A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. Lect. Notes Comput. Sc., 1917: 849-858.

Dias, A.H. and J.A. De Vasconcelos, 2002. Multiobjective genetic algorithms applied to solve optimization problems. IEEE T. Magn., 38(2): 1133-1136.

Dougherty, J., R. Kohavi and M. Sahami, 1995. Supervised and unsupervised discretization of continuous features. Proceeding of the 12th International Conference on Machine Learning, 12: 194-202.

Eddaoudi, F., F. Regragui, A. Mahmoudi and N. Lamouri, 2011. Masses detection using SVM classifier based on textures analysis. Appl. Math. Sci., 5(8): 367-379.

Ghosh, A. and B. Nath, 2004. Multi-objective rule mining using genetic algorithms. Inform. Sciences, 163(1): 123-133.

Gonçalves, J.F., J.J. de Magalhães Mendes and M.G. Resende, 2005. A hybrid genetic algorithm for the job shop scheduling problem. Eur. J. Oper. Res., 167(1): 77-95.

Hall, M.A., 1999. Correlation-based feature selection for machine learning. Ph.D. Thesis, the University of Waikato.

Haralick, R.M., K. Shanmugam and I. Dinstein, 1973. Texture features for image classification. IEEE T. Syst. Man Cyb., 8(6): 610-621.

Heath, M., K. Bowyer, D. Kopans, R. Moore and P. Kegelmeyer, 2000. The digital database for screening mammography. Proceeding of the 5th International Workshop on Digital Mammography, pp: 212-218.

Holte, R.C., 1993. Very simple classification rules perform well on most commonly used datasets. Mach. Learn., 11(1): 63-90.

Ishibuchi, H. and T. Yamamoto, 2004. Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining. Fuzzy Set. Syst., 141(1): 59-88.

Jain, N., V. Sharma and M. Malviya, 2012. Reduction of negative and positive association rule mining and maintain superiority of rule using modified genetic algorithm. Int. J. Adv. Comput. Res., 2(4): 6.

Keshavamurthy, B.N., A.M. Khan and D. Toshniwal, 2013. Privacy preserving association rule mining over distributed databases using genetic algorithm. Neural Comput. Appl., 22(1): 351-364.

Lee, D.G., K.S. Ryu, M. Bashir, J.W. Bae and K.H. Ryu, 2013. Discovering medical knowledge using association rule mining in young adults with acute myocardial infarction. J. Med. Syst., 37(2): 1-10.

Lim, A.H., C.S. Lee and M. Raman, 2012. Hybrid genetic algorithm and association rules for mining workflow best practices. Expert Syst. Appl., 39(12): 10544-10551.

Lisboa, P.J.G., 2002. A review of evidence of health benefits from artificial neural networks in medical intervention. Neural Networks, 15(1):11-39.

Liu, H. and H. Motoda, 2007. Computational Methods of Feature Selection. CRC Press, NY.

Mavroforakis, M.E., H.V. Georgiou, N. Dimitropoulos, D. Cavouras and S. Theodoridis, 2006. Mammographic masses characterization based on localized texture and dataset fractal analysis using linear, neural and support vector machine classifiers. Artif. Intell. Med., 37(2): 145-162.

Nahar, J., T. Imam, K.S. Tickle and Y.P.P. Chen, 2013. Association rule mining to detect factors which contribute to heart disease in males and females. Expert Syst. Appl., 40(4): 1086-1093.

Pang, S., S. Ozawa and N. Kasabov, 2005. Incremental linear discriminant analysis for classification of data streams. IEEE T. Syst. Man Cy. B, 35(5): 905-914.

Park, J. and I.W. Sandberg, 1991. Universal approximation using radial-basis-function networks. Neural Comput., 3(2): 246-257.

Platt, J., 1998. Sequential minimal optimization: A fast algorithm for training support vector machines. Technical Report, MSR-TR-98-14, pp: 21.

Rish, I., 2001. An empirical study of the naive Bayes classifier. IBM Research Report, RC 22230 (W0111-014), New York, pp: 41-46.

Rossquinlan, J., 1993. C4. 5: Programs for Machine Learning. Morgankaufmann Publishers Inc., San Francisco, CA.

Saggar, M., A.K. Agrawal and A. Lad, 2004. Optimization of association rule mining using improved genetic algorithms. Proceeding of the IEEE International Conference on Systems, Man and Cybernetics, 4: 3725-3729.

Shrivastava, V.K., D.P. Kumar and D.K. Pardasani, 2010. Extraction of interesting association rules using GA optimization. Global J. Comput. Sc. Technol., 10(5).

Singh, B.K., 2011. Mammographic image enhancement, classification and retrieval using Color, statistical and spectral analysis. Int. J. Comput. Appl., 27(1).

Soh, L.K. and C. Tsatsoulis, 1999. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE T. Geosci. Remote, 37(2): 780-795.

Tsai, C.F., C.W. Tsai and T. Yang, 2002. A modified multiple-searching method to genetic algorithms for solving traveling salesman problem. Proceeding of the IEEE International Conference on Conference on Systems, Man and Cybernetics, 3: 362-366.

Vafaie, H. and K. De Jong, 1992. Genetic algorithms as a tool for feature selection in machine learning. Proceeding of the 4th International Conference on Tools with Artificial Intelligence (TAI'92), pp: 200-203.

Wakabi-Waiswa, P.P., V. Baryamureeba and K. Sarukesi, 2011. Optimized association rule mining with genetic algorithms. Proceeding of the 7th International Conference on Natural Computation (ICNC, 2011), 2: 1116-1120.

Yu, L. and H. Liu, 2003. Feature selection for high-dimensional data: A fast correlation-based filter solution. Proceeding of the 20th International Conference on Machine Learning (ICML-2003). Washington, DC, 3: 856-863.