

Classification algorithms for microcalcifications in mammograms (Review)

E. SAKKA, A. PRENTZA and D. KOUTSOURIS

Biomedical Engineering Laboratory, National Technical University of Athens,
9 Iroon Polytechniou Str., 15773 Athens, Greece

Received September 6, 2005; Accepted September 28, 2005

Abstract. Early detection is the key to improve breast cancer prognosis. The only proven effective method of breast cancer early detection is mammography. An early sign of 30-50% of breast cancer is the appearance of clusters of fine, granular microcalcifications and 60-80% of breast carcinomas reveal microcalcification clusters upon histological examination. The high correlation between the appearance of the microcalcification clusters and diseases, proves that computer aided diagnosis (CAD) systems for automated classification of microcalcification clusters will be very useful and helpful for breast cancer control. The fuzzy nature of microcalcification, the low contrast and the low ability of distinguishing them from their surroundings make automated characterization of them extremely difficult. In this study, we give an overview of the currently available literature on characterization of malignant and benign microcalcifications. We compare and evaluate some of the classification algorithms on microcalcifications in mammograms used in various CAD systems, which are separated into categories according to the method in use. Neural networks are used in applications where only a few decisions are required concerning an amount of data. The K-nearest neighbour classifier distinguishes unknown patterns based on the similarity to known samples and the decision tree approach is much simpler than neural networks and does not need extensive knowledge of the probability distribution of the features.

Contents

1. Introduction
2. Classification of microcalcifications
3. Discussion

Correspondence to: E. Sakka, Biomedical Engineering Laboratory, National Technical University of Athens, 9 Iroon Polytechniou Str., Computer Building, Office 106, 15773 Zografou Campus, Athens, Greece
E-mail: esakka@biomed.ntua.gr

Key words: microcalcifications, feature extraction, classification, ROC curve

1. Introduction

Breast cancer is the most common type of cancer among women and is the second leading cause of cancer deaths for women today, after lung cancer. According to the World Health Organization, >1.2 million people will be diagnosed with breast cancer this year worldwide. The American Cancer Society estimates that in 2005, approximately 211,240 women in the United States are expected to have been diagnosed with invasive breast cancer (stages I-IV). Another 58,490 women are expected to have been diagnosed with *in situ* breast cancer, a very early stage of the disease (1). Breast cancer is also the most common cancer in females in Europe. It is estimated that in the year 2004 there were 350,000 new breast cancer cases in Europe, while the number of deaths from breast cancer was estimated at 130,000. Breast cancer is responsible for 26.5% of all new cancer cases among women in Europe and 17.5% of cancer deaths (2).

Primary prevention seems impossible since the causes of this disease still remain unknown. Thus, early detection is the key to improve breast cancer prognosis. The only proven effective method of breast cancer early detection is mammography. There are some limitations, though, in human observing and it is difficult for radiologists to evaluate both accurately and uniformly the enormous number of mammograms generated in widespread screening. The presence of microcalcification clusters is an important sign for the detection of early breast carcinoma. High correlation between the appearance of the microcalcification clusters and the disease shows that CAD systems designed for automated classification of microcalcification clusters could be very useful and helpful for breast cancer control.

Although computer-aided mammography has been studied for over two decades, the fuzzy nature of microcalcification, low contrast and low ability of distinguishing them from their surroundings make extremely difficult an automated characterization of microcalcifications and that is because: a) microcalcifications are very small. The sizes of microcalcifications range between 0.1 and 1.0 mm and the average size is about 0.3 mm. Some isolated ones that are <0.1 mm cannot be distinguished in the film-screen mammography due to the high-frequency noise; b) microcalcifications vary in size, shape and distribution; therefore, simple template matching is impossible; c) microcalcifications may have low contrast, resulting in light intensity difference between suspicious areas and their surrounding tissues; d) microcalcifications may be

closely connected to surrounding tissues, in which case simple segmentation algorithms cannot work well; e) in some dense tissues and/or skin thickening, especially in the breast of younger women, suspicious areas are almost invisible. Dense tissues may easily be misinterpreted as microcalcifications and a high false-positive (FP) rate will be yielded, that is a major problem in most of the algorithms. Most FPs are due to the emulsion error, digitization artefacts or anatomical structures such as fibrous strands, breast borders or hypertrophied lobules that look like microcalcifications.

2. Classification of microcalcifications

It is usually very difficult to distinguish between benign and malignant microcalcification clusters because of the variability associated with their appearance. An image region may be called cancerous (positive) or normal (negative) and a decision for a detection result may be either correct (true) or incorrect (false). A decision for a detection result, therefore, will be one of four possible categories: true positive (TP), true negative (TN), false positive (FP) and false negative (FN). FN and FP represent two kinds of errors. An FN error implies that a true abnormality was not detected and an FP error occurs when a normal region was falsely identified as abnormal. A TP decision is a correct judgment of an existing abnormality and a TN decision means that a normal region was correctly labelled. In order to measure and report the performance of a diagnostic system, the following indices are used: sensitivity and specificity. The terms sensitivity and specificity are synonymous with true positive and true negative rate, respectively, that is: $\text{sensitivity} = \text{TPs}/(\text{TPs} + \text{FNs})$ and $\text{specificity} = \text{TNs}/(\text{TNs} + \text{FPs})$. For a fixed discrimination capacity, the sensitivity and specificity of a diagnostic system will depend on the particular confidence threshold that the system uses to partition evidence into categorical decisions. Both sensitivity and specificity will change if the confidence threshold is changed. A solution to this problem is suggested by the use of receiver operating characteristic (ROC) curve which indicates the tradeoffs between sensitivity and specificity that are available from a diagnostic system and thus describes the inherent discrimination capacity of that system. Another way to measure the performance of the classification scheme is with the area under the ROC curve which yields the A_z value. Generally, a larger A_z indicates greater overall performance (3).

Feature extraction. Feature extraction is a very important process for the overall system performance in the classification of microcalcifications. The extracted features should have the following characteristics (4): a) reliability: features should have similar values for all patterns of the same class; b) independence: features should be as much uncorrelated as possible; c) discrimination: features should take on significantly different values for patterns belonging to different classes.

Moreover, the number of features should be sufficiently limited, since the complexity of a CAD system increases rapidly with the number of features used.

The features extracted are distinguished according to the method of extraction and the image characteristics. Thus, the following categorization of the features can be applied: a) features extracted directly from mammogram, such as peri-

meter, area, compactness, elongation, eccentricity, thickness, orientation, direction, line, background, foreground, distance and contrast. They are easy to extract and they originate from the experience of radiologists; b) features extracted from spatial grey level dependence matrix (co-occurrence matrix); c) features extracted from the grey-level run-length (GLRL) matrix; d) features extracted from the grey-level difference (GLD) matrix; e) energy, entropy and norm extracted from the wavelet transform sub-images; f) features extracted from the fractal model of an image; g) features used to describe the distribution of the microcalcification, cluster area and number of microcalcifications in an area. In the following paragraphs we summarise these different methods of feature extraction.

Shape feature extraction. Numerous researchers use shape features extracted from mammogram to directly describe individual microcalcification. Soltanian-Zadeh *et al* (5) select 15 features. These are: number of microcalcifications in cluster, maximum size of calcifications in cluster, standard deviation of the size of calcifications in cluster, number of calcifications with size of one pixel, sum of the area of the calcifications in each cluster, maximum value of compactness (the ratio of the squared perimeter to the area) in cluster, average compactness in cluster, maximum value of the moment which represents the roughness of microcalcifications in cluster, average roughness in cluster, radius of the circle that best fits the cluster, scattering of the microcalcifications, average grey level of the microcalcifications in cluster, standard deviation of the mean of the microcalcification grey levels in cluster, maximum standard deviation of the grey levels in each calcification and average standard deviation of the grey levels in each calcification in cluster.

Veldkamp *et al* (6) use a total amount of 16 features for classification of microcalcification clusters, which are distinguished in two types: distribution features, based on the distribution of individual microcalcification features within a cluster and cluster shape and cluster position features, describing the location of clusters in a mammogram. Distribution features are the number of calcifications in a cluster, the mean and standard deviation of microcalcification area, orientation, contrast, eccentricity and compactness. Cluster shape features are cluster area, cluster eccentricity and cluster orientation. Finally, two cluster position features are investigated: the relative distance between a cluster and the pectoral muscle and the relative distance of a cluster to the breast edge.

Lee *et al* (7) use the shape orientation information and the shape curvature to extract the following features: average size per blob, scattering density of blobs within a ROI, density of microcalcifications, irregularity of microcalcifications, elongation of microcalcifications, circularity of microcalcifications, compactness and defective degree (the zigzag boundary of a round microcalcification).

Sun *et al* (8) use shape features for the classification of microcalcifications such as the ratio of suspicious areas and the ratio of circularity between two ROIs in different views of ipsilateral images.

Foggia *et al* (9) extract four features, based on the analysis of the shape: compactness, roughness, border gradient strength and local contrast.

Lo *et al* (10) extract 22 image processing features. These are the number of calcifications, the logarithm of that number, total area of all calcifications, the logarithm of that area and the mean and standard deviation of each of the following nine morphological features: calcification distance, number of overlaps (resulting from the overlapping ROIs in histogram analysis), calcification area, compactness, central moment, Fourier descriptor, eccentricity, spread and orientation.

Statistical feature extraction. Kim and Park (11) make a comparative study of the performances of the surrounding region dependence method (SRDM) and other conventional statistical texture-analysis methods for detecting clustered microcalcifications in digitized mammograms such as the spatial grey-level dependence method (SGLDM), the grey-level run-length method (GLRLM) and the grey-level difference method (GLDM). Textural features extracted by these methods, such as contrast, angular second moment, entropy, mean and inverse difference moment, are used to classify ROIs into positive ROIs, containing clustered microcalcifications and negative ROIs containing normal tissues.

Kim *et al* (12) use the SRDM to extract the horizontal-weighted sum, the vertical-weighted sum, the diagonal-weighted sum and the grid-weighted sum.

Chan *et al* (13) apply texture analysis for the evaluation of textural changes in breast tissue, derived by a developing malignancy. From the SGLD matrix, they extract the following features: correlation, entropy, energy (angular second moment), inertia, inverse difference moment, sum average, sum entropy and difference entropy, difference average, sum variance, difference variance and information measure of correlation.

Soltanian-Zadeh *et al* (5) extract a set of features using the SGLDM. For quantifying texture, ten features are calculated: angular second moment, contrast, correlation, variance, inverse difference moment, entropy, difference variance, difference entropy and two information measures of correlation.

Instead of applying the co-occurrence matrix method for analyzing the texture variations in the medical images, Kuo *et al* (14) use a statistical method which includes contrast, covariance and dissimilarity to evaluate the texture parameters for several distances between pixels, directly from the image. The main advantage of this method resides in its calculation cost, which depends only on the size of the image treated and not on the number of grey levels.

Halkiotis *et al* (15) use moments of the grey-level histogram of an image or region in order to describe texture. From the central moments, the mean, variance, skewness and kurtosis are estimated for each object. They also estimate the number of objects found in a region of radius of 2 cm around each object.

Sun *et al* (8) extract a series of spatial, morphology and texture features in detected areas. The texture features include the ratio of intensity, the ratio of mean gradient of region boundary between two areas in different image views.

Foggia *et al* (9) use the following textural features for the classification of the clustered microcalcifications: energy in the ROI which is the average square intensity in the ROI; energy in the background which is the average square intensity in the background; average intensity in the ROI; standard deviation of the intensity in the ROI; entropy of the first

order histogram which is a measure of the uniformity of the distribution of the grey levels in the ROI (the higher this parameter is, the more uniform the distribution of grey levels in the ROI gets); energy of the second order histogram which is the average square value of the co-occurrence matrix evaluated on the ROI; contrast of the second order histogram which is a measure of the distribution of the difference among the grey levels exhibited by the points in the ROI and entropy of the second order histogram which is a measure of the uniformity of the distribution of the values in the co-occurrence matrix.

Multiscale texture features extraction - wavelet based method. Wavelet theory provides a powerful framework for multi-resolution analysis and it can be used for texture analysis. By using the multiresolution capability, the wavelet transform could separate small objects such as microcalcifications from large objects such as large background structures.

Bazzani *et al* (16) use the wavelet transform in order to detect signals having size <1 mm by means of a multiresolution analysis.

Ferreira and Borges (17) apply the wavelet transform, using the Daubechies 4 and Haar wavelets, in the decomposition process and keep only the biggest coefficients in magnitude of the decomposed image in first level of decomposition.

Sentelle *et al* (18) employ wavelet analysis to detect calcifications. This is done during wavelet processing by simply removing the lowest resolution approximation coefficients and performing a reverse wavelet transform. The reverse wavelet transform output is then thresholded appropriately to provide binary detections of calcifications.

Soltanian-Zadeh *et al* (5) use three different wavelet packets; Daubechies 6, 10 and 12 wavelets. Their method for feature extraction is based on decomposing the images and calculating the entropy and energy of each of the sub-bands. They also apply the multi-wavelet transformation with GHM, CL and SA4 multi-wavelets, which uses several scaling functions and mother wavelets.

Lambrou *et al* (19) use the Daubechies 4-TAP wavelet filter in all wavelet architectures. From all the signals and their wavelet transformation coefficients, they collect the first and second order statistical values, as well as the grey level run length measurements.

Yoshida *et al* (20) apply the wavelet transformation using the least asymmetric Daubechies' wavelets with length 8, 12 and 20 in order to enhance the microcalcifications.

Cluster features extraction. After individual microcalcifications are detected, cluster features are used to group them into clusters. Veldkamp *et al* (6) investigate a total amount of 16 features for classification of microcalcification clusters. They use two feature types: distribution features (based on the distribution of individual microcalcification features within a cluster) and cluster shape and cluster position features (describing the location of clusters in a mammogram). The distribution features that are used are the number of calcifications in a cluster and the mean and standard deviation of microcalcification area, orientation, contrast, eccentricity and compactness. The cluster shape features that are used are cluster area, cluster eccentricity and cluster orientation.

Finally, they investigate two cluster position features: the relative distance between a cluster and the pectoral muscle and the relative distance of a cluster to the breast edge.

Foggia *et al* (9) extract three types of features; the first concerns the shape of the clusters, the second the distribution of the microcalcifications within the cluster and the third is based upon the presence of microcalcifications within irregular and non-uniform shapes, where the more irregular and the less uniform the shapes of the microcalcifications are, the more likely a malignant cluster is.

Sorantin *et al* (21) use a set of 12 features in order to classify the clusters of microcalcifications in benign and malignant such as the number of microcalcifications in cluster, maximum of microcalcification areas within the same cluster, average of microcalcification areas within the same cluster, standard deviation of microcalcification areas within the same cluster, range of microcalcification areas within the same cluster, standard deviation of microcalcification perimeters within the same cluster, maximum of inter microcalcification distances within the same cluster, standard deviation of inter microcalcification distances within the same cluster, variance of distances between microcalcifications in cluster, minimum of horizontal extension in cluster, cluster area and cluster perimeter.

Classifiers. Classifiers play an important role in the implementation of computer-aided diagnosis of mammography. The features or a subset of these features are employed by classifiers to classify microcalcifications into benign and malignant. A brief report of five kinds of classifiers is given below.

Neural networks. The artificial neural networks (ANNs) are non-parametric pattern recognition systems that can extract general rules by learning from examples. They are particularly useful in problems where decision rules are vague and there is no explicit knowledge about the probability density functions governing sample distributions. The key characteristics of the artificial neural networks are the distributed representation, the local operations and non-linear processing. When needed, the neural networks can be easily updated to learn in new features.

Kim and Park (11) use a three layer backpropagation artificial neural network (BP-ANN), with textural features used as the input signals of the input layer. A comparative study of the classification accuracy is performed for the SRDM, the SGLDM, the GLRLM and the GLDM. The result of the study is that from the viewpoint of classification accuracy and computational complexity, the SRDM is superior to the other conventional methods. The A_z values of the SRDM (with 5 hidden neurons), the SGLDM (with 10 hidden neurons), the GLDM (with 7 hidden neurons) and the GLRLM (with 7 hidden neurons) are 93%, 88%, 74% and 68%, respectively.

Sorantin *et al* (21) implement two different three layer ANNs achieving an A_z of 64%, when training the ANN with all clusters of all patients.

Lo *et al* (10) use a BP-ANN classifier to determine whether each ROI contains a calcification and a second BP-ANN classifier to reduce the number of false positive clusters.

Thus, they result in a very poor outcome, where the best A_z is only 69% and the best partial A_z , which is the average specificity over the range of sensitivities from 90% to 100%, is 21%.

Kim *et al* (12) also employ a three layer BP-ANN as a classifier with four hidden neurons and obtain a sensitivity of >90% with a low FP detection rate of 67% per ROI.

Chan *et al* (13) train a BP-ANN classifier and test it with the leave-one-case-out method to recognize the malignant or benign microcalcification clusters. The performance of the ANN is analysed and it is found that with a subset of six texture features out of 22 the ANN classifier achieves an area under the ROC curve of 88%.

Verma and Zakos (22) employ a number of feature extraction techniques in order to extract a combination of 3 features (entropy, standard deviation and number of pixels) and find that this is the best combination to distinguish a benign microcalcification pattern from one that is malignant. Then, they use a feedforward BP-ANN for classification. The BP-ANN architecture uses only one hidden layer. The classification was implemented in the Nijmegen mammographic database and the achieved TP is 88.9%.

Lee *et al* (7) use a shape recognition-based ANN, built with universal feature planes, called Shape Cognitron (S-Cognitron). The system is evaluated by using the Nijmegen mammogram database and the experimental results show that sensitivity and specificity can reach 86.1% and 74.1%, respectively.

Bocchi *et al* (23) implement the classification algorithm using a Kohonen layer followed by a multilayer feed-forward network. The use of one hidden layer in the neural network give 91% TP, whereas without using a hidden layer the result was 82% TP.

Halkiotis *et al* (15) invoke a multi-layer perceptron and evaluate the performance, using two different topologies, with five and ten hidden nodes respectively. The system with five hidden neurons achieves a sensitivity of 84.2% with an average of 26% FP findings per image, whereas the second one succeeds a sensitivity of 94.7% with an average of 27% FP findings per image.

Grohman and Dhawan (24) use a convex-set based neuro-fuzzy algorithm for the classification of 191 difficult-to-diagnose instances of breast cancer. A 20-dimensional set of difficult-to-diagnose mammographic microcalcifications is used to evaluate the neuro-fuzzy pattern classifier (NFPC) and the backpropagation methods. The comparative results show that the ability of the convex-set based method to infer knowledge was better than that of backpropagation in all the performed tests, making it more suitable for use in real diagnostic systems. The maximum mean ROC area that is achieved when using 90% of the features, is 88.37%.

Sun *et al* (8) develop a supervised three-layer ANN classifier, in which backpropagation algorithm, combined with Kalman filtering is used in the neural network's training. Due to the Kalman filter gain introduced to the BP algorithm, the adaptive convergence rate of the algorithm is achieved. Thus, the Kalman filter based backpropagation algorithm exhibits fast convergence, based on a generalized delta learning rule. The system achieves 90% sensitivity with 3 FP findings.

Cheng *et al* (25) use a three-layer, feed-forward, error BP-ANN in order to classify the detected clusters of microcalcifications using as inputs the mean and the standard deviation and achieves an accuracy >97% TP rate with the FP rate of three clusters per image. The mammographic database they use is the Nijmegen database.

Lauria *et al* (26) perform the microcalcification analysis using two different ANN. The first one is a Feed Forward Neural Network that classifies windows in which images are segmented. The output of this ANN becomes the input for the second ANN which by using the principal component method, classifies ROIs. In the classification of clusters the system gets a sensitivity of 92% and a specificity of 92%.

Sajda *et al* (27) implement a Hierarchical Pyramid Neural Network (HPNN), where the ANN are multilayer perceptrons, having one hidden layer with between four and eight hidden units. Each network in the HPNN hierarchy receives input from the integrated feature pyramid and the input for the hidden unit from networks lower in the hierarchy. Networks are trained either coarse-to-fine or fine-to-coarse, depending on the architecture. In the coarse-to-fine HPNN, the lowest network in the hierarchy is first trained until convergence and then all parameters in this network are held fixed, while the next network on the hierarchy is trained. Coarse-to-fine training is possible because the positions of the small objects are well defined when the resolution is decreased. For the fine-to-coarse HPNN, extended objects do not have a definite location at high resolution. The area of ROC curve is 94% while the FP findings are 21%.

Foggia *et al* (9) implement an expert system in order to classify a cluster containing microcalcifications. Each microcalcification is classified by the mC-Expert, while the cluster, considered as a whole, is classified by the Cluster Expert. The final classification decision is obtained by collecting their responses and applying a suitable decision scheme, based on the evaluation of the reliability of each classification. For both classification systems, BP-ANN is used. The architecture of the ANN consists of three layers, containing 25 neurons in the hidden layer, 2 output neurons (associated with the benign and the malignant class) and a number of nodes in the input layer depending on the size of the feature vector employed. The system is evaluated by the ROC curve and achieves 89% sensitivity and 70% FP findings.

K-nearest neighbour classifiers. K-nearest neighbour (KNN) classifier distinguishes unknown patterns based on the similarity to known samples. The KNN algorithm computes the distances from an unknown pattern to every sample and selects the K-nearest samples as the base for classification. The unknown pattern is assigned to the class containing the most samples among the K-nearest samples.

Soltanian-Zadeh *et al* (5) use a KNN classifier and the decision criterion they choose is a measure of malignancy. The input features are extracted using the multi-wavelet method and the area in the ROC curve reaches 89% for CL multi-wavelet. This approach generates TP findings of 85% with FP findings of 10%.

Veldkamp *et al* (6) use a classification method consisting of two classification steps. The first step classifier is used for classification of clusters based on the KNN method. As a

result of this classification step, each cluster is assigned a likelihood. For the final patient-based classification result, the classifier combines classification results of corresponding clusters in medio-lateral oblique and craniocaudal views by taking their mean likelihood of malignancy. The system generates A_z 83%.

Nearest neighbour classifiers based on Euclidean distance. Ferreira and Borges (17) design a nearest neighbour classifier, for the verification of the classification, using Euclidean distance as a metric between the correspondent wavelet coefficients.

Sentelle *et al* (18) calculate the Euclidean distance in order to classify the calcifications and the TP is nearly 80%.

Binary decision tree. Decision tree is a classification method that generates a tree to classify a set of input examples according to their class. Each branch in the tree represents a decision. Each node in the tree refers to a particular attribute. Edges connecting nodes are labelled with attribute values and leaf nodes give a classification that applies to the examples that were reached through that branch. At each step of the tree construction a node is selected according to a statistical measure called information gain, that measures how well a node (attribute) distributes the input examples with respect to their class.

Kuo *et al* (14) use data mining with a decision tree model to classify breast tumours. The co-variance texture parameters are used as the input to construct the decision tree model. This method results in 94.6% sensitivity and 85.4% specificity.

Flores and Gonzalez (28) implement a decision tree in the data mining task, in order to generate descriptive patterns based on a set of characteristics. In this decision tree the root node is the area node. If area has a value of ≤ 13 , the value of the diameter attribute is verified. If diameter has a value of ≤ 2 , the class of the example is negative. If the value for diameter is > 2 , the value of the diameter attribute is verified again and if it is ≤ 2.83 , the convexity attribute is verified. If the convexity attribute has a value of ≤ 0.93 , the class is negative, otherwise the class is positive. The achieved TP with decision trees is 92.6%.

Support vector machines. Support vector machines (SVM) are learning machines used in pattern recognition and regression estimation problems. They grow up from statistical learning theory (SLT) problems, which give some useful bounds on the generalization capacity of machines for learning tasks. The SVM algorithm constructs a separating hypersurface in the input space. It acts in this way: a) it maps the input space into a high dimensional feature space through some non-linear mapping chosen a priori (Kernel); and b) constructs the maximal margin hyperplane in this feature space.

Bazzani *et al* (16) construct an SVM classifier and achieves a sensitivity of 95% with 60% FP clusters per image on the 40 images of the Nijmegen database.

3. Discussion

Automated classification of microcalcifications in breast has been an issue of investigation for many researchers over the

Table I. An overview of classifiers for microcalcifications classification.

Classifier	Data set	Feature extraction	Performance	Refs.
Neural networks	172 ROI's were selected from 120 digitized mammograms, source N/A	SRDM	A_z : 93%	11
		SGLDM	A_z : 88%	
		GLDM	A_z : 74%	
		GLRLM	A_z : 68%	
	272 mammographic films, source N/A	Cluster features	A_z : 64%	21
	292 cases, digital database for screening mammography	Features extracted directly from the mammogram and morphological features	A_z : 69%	10
	120 mammograms, 140 ROIs, database from Department of Radiology at Asan Medical Center in Korea	SRDM	Sensitivity 90% FP 67% per ROI	12
	86 mammograms, Department of Radiology at the University of Michigan	SGLD	A_z : 88%	13
	18 microcalcification areas, database of digital mammograms from University of Nijmegen (The Netherlands)	Entropy, standard deviation and number of pixels	TP: 88.9%	22
	40 mammograms, 102 ROIs, database of digital mammograms from University of Nijmegen (The Netherlands)	Features extracted directly from the mammogram	Sensitivity: 86.1% Specificity: 74.1%	7
	30 mammograms, 50 ROIs, source N/A	Morphological features	TP: 91%	23
	25 mammograms, database by the Mammographic Image Analysis Society (MIAS)	Textural features	Sensitivity: 84.2% FP: 26% (5 hidden neurons) Sensitivity: 94.7% FP: 27% (10 hidden neurons)	15
	191 difficult-to-diagnose mammograms	Wavelet features	A_z : 88.37%	24
	40 mammograms, source N/A	Shape features	Sensitivity: 90%	8
	40 mammograms (Nijmegen database)	Statistical features	TP: 97% FP: 3 clusters per image	25
K-nearest neighbour classifier	806 microcalcification clusters, source N/A		Sensitivity: 92% Specificity: 92%	26
	136 ROIs, source N/A	Maximum energy, pixel location, energy of the orientation	A_z : 94% FP: 21%	27
	40 mammograms (Nijmegen database)	Shape features	Sensitivity: 89% FP: 70%	9
	40 mammograms (Nijmegen database)	Wavelet and multi-wavelet features	A_z : 89% TP: 85% FP: 10%	5
	165 mammograms Nijmegen database	Distribution features cluster shape and cluster position features	A_z : 83%	6
Nearest neighbour classifiers based on Euclidean distance	25 images, digital database for screening mammography (DDSM)	Wavelet features	TP: 80%	18
Binary decision tree	259 mammograms, HDI 3000 database	Statistical features	Sensitivity: 94.6% Specificity: 85.4%	14
	84 mammograms, database of ISSSTEP hospital	Features extracted directly from the mammogram	TP: 92.6%	28
Support vector machines	40 mammograms (Nijmegen database)	Wavelet features	Sensitivity: 95% FP: 60%	16

last 20 years. Although great steps towards the amelioration of these techniques have been made, there are still remaining challenges and directions for future research; first in the field of designing better feature detection and selection algorithms, in order to distinguish in a more efficient way between the malignant and the benign microcalcification clusters; secondly in the integration of classifiers to reduce both FPs and FNs. In order to reduce false positive, several different types of features, sometimes, with clinical information should be used. Obviously the higher the complexity of the algorithm is, the more time-consuming the algorithm gets. Thus it is important for the researchers to implement algorithms balanced between accuracy and computational complexity. Another important issue is to use a standard database for the evaluation of the algorithms.

According to the comparative results shown in Table I, it is obvious that the classifiers based on artificial neural networks is a method that can result in higher detection rate and lower false positive findings in clustered microcalcifications. Moreover, the feature vector that is used plays an important role to the classification rates. The implementation of statistical methods for feature extraction reveals higher A_z values in the classification process.

References

1. American Cancer Society: Breast Cancer Facts and Figures. <http://www.cancer.org/>
2. Breast cancers. In: Health Statistics-Atlas on Mortality in the European Union. Eurostat, http://epp.eurostat.cec.eu.int/cache/ITY_OFFPUB/ATLAS/EN/ATLAS-EN.HTM
3. Woods K, Sallam MY and Bowyer KW: Evaluating detection algorithms. In: Digital Mammography Nijmegen 98. Karssemeijer N, Thijssen M, Hendriks J and van Erning L (eds). Kluwer Academic Publisher, Amsterdam, pp19-45, 1998.
4. Castleman KR: Digital Image Processing. Prentice-Hall, Englewood Cliffs, NJ, pp527-534, 1996.
5. Soltanian-Zadeh H, Rafiee-Rad F and Pourabdollah-Nejad DS: Comparison of multiwavelet, wavelet, Haralick and shape features for microcalcification classification in mammograms. *Pattern Recognition* 37: 1973-1986, 2004.
6. Veldkamp WJH, Karssemeijer N and Hendriks JHCL: Experiments with radiologists and a fully automated method for characterization of microcalcification clusters. *International Congress Series* 1230: 586-592, 2001.
7. Lee SK, Chung PC, Chang CI, Lo CS, Lee T, Hsu GC and Yang CW: Classification of clustered microcalcifications using a Shape Cognitron neural network. *Neural Networks* 16: 121-132, 2003.
8. Sun X, Qian W and Song D: Ipsilateral-mammogram computer-aided detection of breast cancer. *Comput Med Imaging Graph* 28: 151-158, 2004.
9. Foggia P, Sansone C, Tortorella F and Vento M: Automatic classification of clustered microcalcifications by a multiple classifier system. *Multiple Classifier Systems* 2001: 208-217, 2001.
10. Lo JY, Gavrielides M, Markey MK and Jesneck JL: Computer-aided classification of breast microcalcification clusters: merging of features from image processing and radiologists. *Medical Imaging 2003: Image Processing, Proceedings of the SPIE* 5032: 882-889, 2003.
11. Kim JK and Park HW: Statistical textural features for detection of microcalcifications in digitized mammograms. *IEEE Trans Med Imaging* 18: 231-238, 1999.
12. Kim JK, Park JM, Song KS and Park HW: Detection of clustered microcalcifications on mammograms using surrounding region dependence method and artificial neural network. *VLSIVideo* 18: 251-262, 1998.
13. Chan HP, Sahiner B, Petrick N, Helvie MA, Lam KL, Adler DD and Goodsitt MM: Computerized classification of malignant and benign microcalcifications on mammograms: texture analysis using an artificial neural network. *Phys Med Biol* 42: 549-567, 1997.
14. Kuo WJ, Chang RF, Moon WK, Lee CC and Chen DR: Computer-aided diagnosis of breast tumors with different US system. *Acad Radiol* 9: 793-799, 2002.
15. Halkiotis S, Mantas J and Botsis T: Computer-aided detection of clustered microcalcifications in digital mammograms. *Res Systemica Eur Systems Sci J* 2, 2002.
16. Bazzani A, Bevilacqua A, Bollini D, Campanini R, Lanconelli N, Riccardi A and Romani D: Automatic detection of clustered microcalcifications using a combined method and an SVM classifier. *ESANN'2000 proceedings - European Symposium on Artificial Neural Networks, Belgium*, pp195-200, 2000.
17. Ferreira CBR and Borges DL: Analysis of mammogram classification using a wavelet transform decomposition. *Pattern Recognition Lett* 24: 973-982, 2003.
18. Sentelle S, Sentelle C and Sutton MA: Multiresolution-based segmentation of calcifications for the early detection of breast cancer. *Real-Time Imaging* 8: 237-252, 2002.
19. Lambrou T, Linney AD, Speller RD and Todd-Pokropek A: Statistical Classification of Digital Mammograms Using Features from the Spatial and Wavelet Domains. *Medical Image Understanding and Analysis, Portsmouth*, pp22-23, 2002.
20. Yoshida H, Doi K, Nishikawa RM, Muto K and Tsuda M: Application of the wavelet transform to automated detection of clustered microcalcifications in digital mammograms. *Acad Reports, Tokyo Inst Polytech* 17: 24-37, 1994.
21. Sorantin E, Schmidt F, Mayer H, Becker M, Szepesvari C, Graif E and Winkler P: Computer aided diagnosis of clustered microcalcifications using artificial neural nets. *J Comput Information Technol* 2: 151-160, 2000.
22. Verma B and Zakos J: A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques. *IEEE Trans Inf Technol Biomed* 5: 46-54, 2001.
23. Bocchi L, Coppini G, Nori J and Valli G: Detection of single and clustered microcalcifications in mammograms using fractals models and neural networks. *Med Eng Phys* 26: 303-312, 2004.
24. Grohman WM and Dhawan AP: Fuzzy convex set-based pattern classification for analysis of mammographic microcalcifications. *Pattern Recognition* 34: 1469-1482, 2001.
25. Cheng HD, Wang J and Shi X: Microcalcification detection using fuzzy logic and scale space approaches. *Pattern Recognition* 37: 363-375, 2004.
26. Lauria A, Palmiero R, Forni G, Cerello P, Golosio B, Fauci F, Magro R, Raso G, Tangaro S and Indovina PL: The CALMA system: an artificial neural network method for detecting masses and microcalcifications in digitized mammograms. *Nucl Instruments Methods Phys Res A* 518: 391-393, 2004.
27. Sajda P, Spence C and Pearson J: Learning contextual relationships in mammograms using a hierarchical pyramid neural network. *IEEE Trans Med Imaging* 21: 2002.
28. Flores BA and Gonzalez JA: Data mining with decision trees and neural networks for calcification detection in mammograms. *MICAI 2004, LNAI* 2972: 232-241, 2004.