International Journal of Engineering & Technology, 7 (2.25) (2018) 89-94

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper



A cad system for improving classification performance in breast cancer detection

Bincy Babu [1] *, A. Josephin Arockia Dhivya [1], R. Chandrasekaran [1], T. R. Thamizhvani [1], R.J.Hemalatha [1]

[1] *Department of Biomedical Engineering, Vels Institute of Science, Technology and Advanced Studies, Pallavaram, Chennai-600117, India*

**Corresponding author E-mail:* a.dhivya.se@velsuniv.ac.in

Abstract

Early detection is a key factor in reducing breast cancer mortality rate. Research works in the area of mammography plays an important role in identification of calcification clusters and detection of breast cancer. The purpose of proposed research is to find the best combination of feature extraction algorithm to classify mammogram into benign and malignant. It includes Marker Controlled Watershed Segmentation Technique (MCWS), feature set extraction methods and SVM classifier algorithm. The GLCM, GLRLM and first order texture descriptors are used to describe the calcification clusters. The standard inputs such as normal and abnormal breast images for the proposed system are taken from Digital Database for Screening Mammography (DDSM). The computational study showed that combination of all the three features descriptors provide better classification result with 97% accuracy and it ensures improved the CAD system performance for small training data sets compared to existing techniques.

Keywords: *Micro Calcification; Marker Controlled Watershed Segmentation; GLCM; GLRLM.*

1. Introduction

According to World Health Organization (WHO) survey, breast cancer is a prominent cause of death worldwide, accounting for 571 000 deaths in 2015 and ranked as first common cause of cancer death among Indian females with mortality 12.7 per 100,000 women. It continues as important public health issue among women. Early detection is the key factor to reduce cancer death. A better treatment result can be obtained if breast abnormality is identified at the initial stage [1].

Most reliable and commonly used imaging system for detecting breast carcinoma is mammography. However, difficulties come in providing accurate results which can be solved by the proposed Computer Aided Diagnosis (CAD) system. Breast calcifications are calcium deposits in breast tissues and micro calcifications are significant sign that leading to cancer detection [2]. The correlation between disease and calcification cluster appearance proves the necessity of CAD systems in mammogram analysis. Micro calcifications have an area in range of 0.1 to 1 mm and it is very challenging to isolating these calcifications. The size, distribution and shape may vary depending up on the stage of breast cancer that leads to failure of standard CAD detection system which works based on single model. Mammograms are low contrast images, considering the difference in intensity of micro calcification and background is complicated. The chances for misinterpretation of dense tissue mammogram are more compared to glandular and fatty tissues. All these problems can be eliminated by suppressing the noise present in the mammogram and later enhancing the contrast of image. The basic steps involved in any CAD system are preprocessing, segmentation, feature extraction and classification. Segmentation methods identify Region of Interest (ROI) which is varying for different application areas [3].

Different feature extraction methods are used in image processing. Dana cobaz et.al [3] proposed a multi scale Gabor type feature set to obtain accurate result with less number of mammograms in training set. In Albert law et.al [4] studies, deformable region model and sampling techniques are used, in which processing time is reduced and is efficient in timely decision making. The presence low contrast background features can sometimes influence in decision making, it can be extracted by applying wavelet transform and wavelet packets as explained by Azadeh et al [5]. Micro calcification features such as compactness, thickness, direction, foreground, and contrast are extracted in studies of K.S. Woods et al [6] and B. Zheng [7]. Elongation eccentricity, orientation and foreground features are explained and analyzed in mammogram by Veldkamp et al [8] and Bothorel et al [9]. The individual features are easy to extract and are devised by experienced radiologists. Ball et al [10] used micro calcification features that comprise morphological and statistical features constructed on the segmented boundary and with application of rubber band straightening transform.

Timp et al [11] suggested temporal single view features in which these features are categorized into 12 groups based on type of nature. Fauci et al [12] extracted features that gave information related to shape and geometry. The benchmark for feature selection was created by the difference between healthy and pathological regions. Rangayyan et al [13] proposed a method to obtain shape features from the turning angle functions of contours. Features are useful in the analysis of contours of breast masses and tumors because of their ability to capture diagnostically important details of shape related to spicules and lobules. In L. Zhang et al [14] works, cluster features are taken to define the number, area and scattering of the micro calcification.

The major part of a CAD system is classifiers. Based on the extracted feature set classifiers perform classification to identify malignant and benign mammograms. Image processing comprises

Copyright © 2018 Bincy Babu et. al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

many classification techniques which give promising results in classification; they are Random Forest (RF), Multi-Layer Perceptron (MLP), Artificial Neural Network (ANN) etc [15]. Komal Sharma et al [16] proposed a MLP classifier based mammogram analysis approach which had a drawback that the local minima affect the training progression. RF classifier is applied in mammogram by Blagojce et al [17] which is accomplished by directing a sample in every tree followed by allocating it in the marker of the terminal node. The classification result is the mean vote of altogether trees. This system is efficient with high dimensional data and large dataset but has limitation that it may over fit sometimes.

X Wu et al [18] suggested K Nearest Neighbor (KNN) classifier, simple and commonly used where classification is based on majority. Even though it is easy to understand and implement, KNN is computationally expensive for huge set of images. In Naive Bayesian (NB) system for micro calcification detection by R.K.Kavitha et al [19], the limitation of KNN can be avoided and can be implemented for large datasets. Complicated iterative parameters are not needed here. It originates conditional probability from the correlation between each instant class and feature. In Krishnaveni et al [20] works, NB classifier needs less training samples but have less accuracy. Relevance Vector Machine (RVM) classifier networks are used in L Wei et al [21] works which is framed on Bayesian theory. A two stage classification network is adapted for improving the computational speed. It involves the application of linear RVM classifier to reject the overpowering majority. In particular, it is beneficial for real-time analysis because time required for detection is less. Drawback of this system comes is that training stage is complex.

In Vijayarani et al [22] study of breast cancer detection, Decision Table (DT) technique is used. Two different decision tables are used, they are decision table majority and decision table local. If decision table cell that mapping to new occurrence is empty, then majority training sets are returned in majority decision table. This system provides classification errors for small training sets and it necessities more calculations for composite problems. Back propagation algorithm for classification is proposed by Dheeba j et al [23] needs more training time in which the generated error by the difference in target and actual outputs depends on weight of neurons in network layers. The limitations noticed from existing methods are miscalculation due to absence of texture data, over fitting of omitted data, requirement of more training examples, high computation time and trouble to detect complex of calcification. These problems can be overcome by the proposed CAD system. It comprises of feature extraction and classification of mammogram to analyze nature of calcification which further leads to accurate detection of breast cancer.

2. Materials and methods

In this proposed approach, 200 DDSM images are considered for analysis in which 100 mammograms are for training and 100 are for testing. The training dataset included 50 images for benign cases and 50 for malignant. Fig.1 shows framework of proposed system which comprises of segmentation, feature extraction and classification. Each section in detail is explained below.

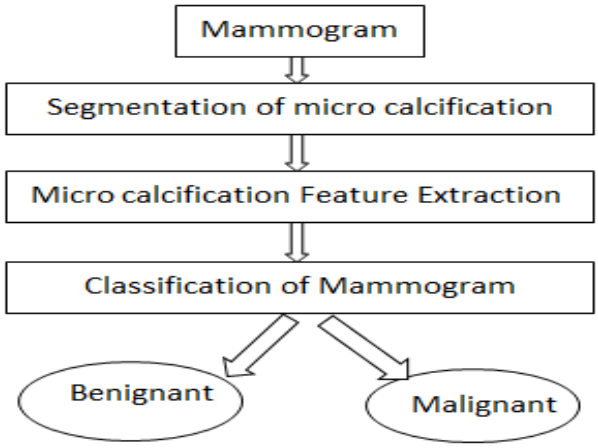


Fig. 1: Framework for Proposed System.

2.1. Mammogram

The mammogram images for proposed study are taken from Digital Database for Screening Mammography (DDSM) which is publically available database for research related purpose of breast cancer and it comprises almost 2500 mammograms [24]. It provides information of normal, benign and cancer mammogram cases. Each case in DDSM is detailed with not only rating of breast density but also keyword explanation of American College of Radiologist (ACR).

2.2. Segmentation of micro calcification

Image preprocessing becomes essential in removing background noises and markings [25]. Mammogram preprocessing includes filtering, contrast enhancement, label and pectoral muscle removal [26]. In the suggested system, noises like salt and pepper are removed by applying adaptive median filter [27] and contrast is improved by Contrast Limited Adaptive Histogram Equalization (CLAHE) [28] technique. The morphological reconstruction techniques are applied to eliminate pectoral muscle and mammogram label. The micro calcification is then segmented by Marker Controlled Watershed Segmentation approach. The resulting Region Of Interest (ROI) showed micro calcification cluster from which features can be extracted.

2.3. Micro calcification feature extraction

The feature extraction technique is implemented in this CAD system to characterize particular properties of the breast calcifications for classification purpose. Mammogram texture feature sets such as entropy, regularity, energy and entropy are extracted from ROI which gave details regarding intensities and spatial arrangement of pixels. Different feature extraction algorithms adapted in this study are explained below.

2.3.1. First order statistic feature

First Order Statistic (FOS) is created in relation with statistical properties of histogram intensity [26]. The features taken from this approach are standard deviation, third moment, smoothness, entropy, uniformity and mean. Here N indicates probable intensity levels, Intensity is indicated by x and intensity histogram by p(x). All features based on first order are expressed as shown in equations (1-6).

$$\text{Entropy (e)} = \sum_{i=0}^{L-1} p(x_i)\log_2 p(x_i) \quad (1)$$

$$\text{Standard Deviation} = \sqrt{\sum_{i=0}^{N-1}(X_i - m)^2\, p(x_i)} \quad (2)$$

$$\text{Uniformity (U)} = \sum_{i=0}^{N-1} p^2(x_i) \quad (3)$$

$$\text{Third Moment } (\mu_3) = \sum_{i=0}^{N-1}(X_i - m)^3\, p(x_i) \quad (4)$$

$$\text{Smoothness } (R) = 1 - \frac{1}{1+\sigma^2} \quad (5)$$

$$\text{Mean} = \sum_{i=0}^{N-1} xp(x_i) \quad (6)$$

2.3.2. GLCM

A second order statistical method is applied in this system to extract texture features. In Gray Level Co-occurrence Matrix (GLCM) calculated mean spatial correlation between grey levels in ROI. The spatial matrix is a frequency formulation or combination of brightness values of pixels. GLCM computed in such a way that a re-quantized image (H) is produced from original image. M_g is the number of grey levels that have value either 16 or 32. By perusing intensity of each pixel and neighboring pixel, GLCM is calculated from H. Here angle, θ is restricted to 0°, 45°, 90°, 135° and displacement (l) has values 1, 2, 3…n. Then GLCM is expressed as G (i, j) as shown in equation (7), where $l_x = l \cos\theta$, $l_y = l \sin\theta$.

$$G(i, j) = \sum_x \sum_y P(l_x\, l_y) \quad (7)$$

The Fig.2 represents the GLCM formation of the grey-level image at unit distance and 0°. The boxes in Fig.2 demonstrate pixel intensities 0 and neighboring pixel 1. The number of such pixels is counted and forming GLCM matrix as shown in Fig.2 (b).

(A)

0	0	1	1	1
0	0	1	1	1
0	2	2	2	2
2	2	3	3	3
2	2	3	3	3

(B)

	0	1	2	3
0	2	2	1	0
1	0	4	0	0
2	0	0	5	2
3	0	0	0	4

Fig. 2: Formation of GLCM (A) Example of an Image with 4 Grey Level Images (B) GLCM for Distance 1 and Direction 0°.

In the same way, counting combination of each pixels values are done. As a result, image is transformed into GLCM matrix. The proposed system used contrast, homogeneity and energy features. Usually contrast is considered as local deviations existing in an image and it can define as degree of the level to which an objective is distinct from its background. It is calculated from the image as per the equation (8). Energy terms orderliness of an image which can be calculated as equation (9). Homogeneity is termed as closeness of scattering elements in GLCM as expressed in equation (10).

$$\text{Contrast} = \sum_{i,j}(i-j)^2\, P(i,j) \quad (8)$$

$$\text{Energy} = \sum_{i,j} P(i,j)^2 \quad (9)$$

$$\text{Homogeneity} = \sum_{i,j} \frac{1}{1+(i-j)^2}\, P(i,j) \quad (10)$$

2.3.3. GLRLM

Grey Level Run Length Matrix (GLRLM) is one feature extraction method in which run length is number of neighboring pixels that have equal grey intensity values in certain directions. Fig.3 represents GLRL matrix. GLRLM texture features used here are Long Runs Emphasis (LRE), Run Length Non-uniformity (RLN) and Shot Runs Emphasis (SRE) [27]. The mathematical expressions for GLCM feature are shown in equations (11-14) where p is number of grey levels, q is length of longest run, i and j indicates length. MATLAB Image processing tools are used to calculate these features.

(A)

1	2	3	4
1	3	4	4
3	2	2	2
4	1	4	1

(B)

Gray level	Run length (j)			
i	1	2	3	4
1	4	0	0	0
2	1	0	1	0
3	3	0	0	0
4	3	1	0	0

Fig. 3: Formation of GLRLM (A) Matrix of Image 4X4 Pixels (B) GLRL Matrix.

LRE processes the spreading of long runs which is predicted big for uneven structural textures. Resemblance of length of runs is expressed as RLN. If run lengths are identical throughout the image, then value of RLN is small.

$$SRE = \frac{1}{n_r}\sum_{i=1}^{R}\sum_{i=1}^{S}\frac{Q[i,j]}{j^2} \quad (11)$$

$$n_r = \sum_{i=1}^{R}\sum_{i=1}^{S} Q[i,j] \quad (12)$$

$$LRE = \frac{1}{n_r}\sum_{i=1}^{R}\sum_{i=1}^{S} Q[i,j] * j^2 \quad (13)$$

$$RLN = \frac{1}{n_r}\sum_{i=1}^{R}\left(\sum_{i=1}^{S} Q[i,j]\right)^2 \quad (14)$$

2.3.4. Classification of mammogram

The suggested method used Support Vector Machine (SVM) classifier for detection of breast cancer in which texture feature is a key descriptor for all mammograms. Texture information is extracted from segmented image and given to proposed classifier. SVM has several benefits compared to current methods; they are higher efficiency and speed in analyzing breast calcification as malignant or benignant. It accomplished low generalization error when applied to categorize mammograms that were not comprised in training set. SVM deliberated as dynamic part of research related to machine learning. It used a premise space of linear function and is trained with an algorithm, based on theory of optimization. The devising practices Empirical Risk Minimization (ERM) principle and Structural Risk Minimization (SRM) principle. SRM underestimates superior bound on the predictable risk and ERM reduces training data errors. It equips SVM as a good classifier. The basic concept of an SVM classifier is illustrated in Fig.4, in which data vectors X and

O are separated by a hyper plane. Many separating planes are considered but SVM pursues a plane that creates biggest margin between data vectors and achieve SRM.

Most probably data points are linearly separable, if not nonlinear transformation is applied for mapping data vectors into feature space. SVM practices a kernel function to eliminate problems related to over fitting in feature space. Equation.15 shows mathematical expression for SVM classifier. Here $\emptyset(x)$ is transformation function and α represents constant.

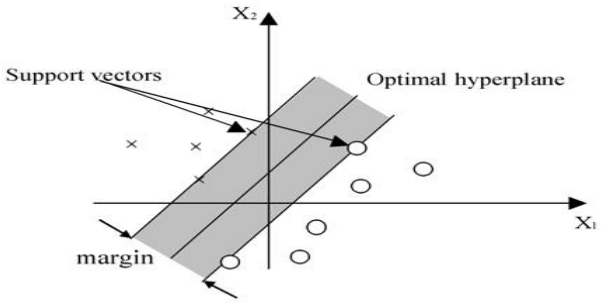


Fig. 4: SVM Classification with a Hyperactive Plane.

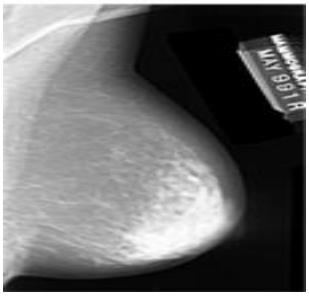
In SVM, support vectors are fundamentals of training set that comprised of samples that are challenging in classification. The components may be on or inside the decision boundary. The kernel function have key role in obtaining better classification result.

$$h(x) = \sum_{i \in SV} \alpha \emptyset(X_i)^T \emptyset(X) + s \quad (15)$$

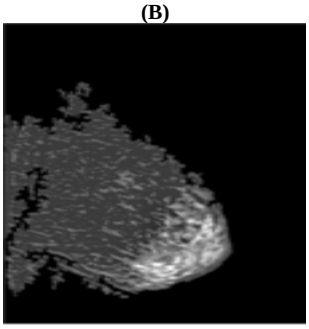
3. Results and discussion

For automatic detection of breast cancer, more than 200 images (100 for testing and 100 for raining) are taken from DDSM database. Mammograms are filtered and enhanced; followed by segmentation to get ROI. Different micro calcification features are extracted from segmented image. The Fig.5 shows outcomes of mammogram segmentation.

(A)



(B)



(C)

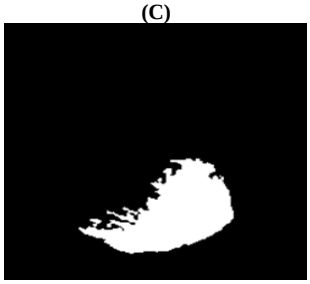


Fig. 5: Mammogram ROI Extraction (A) Original Mammogram (B) Pre Processing Result (C) Segmented Micro Calcification.

The training and testing datasets were used for SVM classification. The training data sets included two cases, benignant and malignant. The first order static features, GLCM and GLRLM features are taken for 100 segmented DDSM mammograms; the minimum and maximum values of benignant and malignant mammogram are tabulated for reference as shown in Table.1. The feature values were different for each case. Mean, smoothness, uniformity, standard deviation, entropy and third moment are attained from first order techniques. Contrast, homogeneity and energy are calculated from GLCM. SRE, RLN and LRE are obtained from GLRLM.

Table.1: Micro Calcification Feature Values for Benign and Malignant Mammograms

Feature extraction technique	Features	Benign Mammogram		Malignant Mammogram	
		Min	Max	Min	Max
First Order Statistic Feature	Mean	54.234	79.581	18.204	47.298
	Smoothness	0.0101	0.0199	0.0214	0.0359
	Standard Deviation	10.968	15.316	24.181	36.017
	Uniformity	0.0302	0.0505	0.0604	0.1872
	Third Moment	-0.0770	0.2628	-0.2168	0.7844
	Entropy	2.086	3.1825	0.7562	1.9214
GLCM	Contrast	0.2541	0.5731	0.0147	0.0829
	Energy	0.33589	0.7571	0.8784	0.9733
	Homogeneity	0.9013	0.9827	0.9912	0.9925
GLRLM	SRE	0.0674	0.0889	0.01295	0.03894
	LRE	5864.778	9962.14	990.233	1526.98
	RLN	32.830	523245	12.337	15.709

In order to analyze performance of classifier, different texture features were elected as input to SVM. SVM classify testing data by training all these feature values. After the training process, the testing stage starts to process. A structure argument generated from training data compared with features calculated from testing data. Table.2 shows confusion matrix for SVM testing with GLCM features. The True Positive (TP) represents number of malignant images that are properly recognized; False Positive (FP) indicates number of benignant images that are identified mistakenly as malignant.

Table.2: Confusion Matrix for SVM Classifier with GLCM Features.

Actual	Predicted	
	Positive	Negative
Positive	True Positive (Tp)	False Positive (Fp)
25	23	2
Negative	False Negative (Fn)	True Negative (Tn)
25	2	23

Table.3: Confusion Matrix for SVM Classifier with GLRLM Features.

Actual	predicted	
	positive	negative
Positive	true positive (tp)	false positive (fp)
25	21	4
negative	false negative (fn)	true negative (tn)
25	3	22

The False Negative (FN) indicates number of malignant image that are not detected. True Negative (TN) designates number of benign mammograms which are correctly detected as benign. The confusion matrix for SVM classification based on features extracted from GLRLM is tabulated in table Table.3. Here four non-cancerous images were detected as cancerous and two cancerous images misclassified as normal. The number of TP and TN in GLCM based SVM classification are more than GLRLM based classification which means GLCM outperforms GLRLM. Table.4 implies confusion matrix of SVM for combination of GLRLM, GLCM and first order statistic features. Number of false negative was same as that of classifier with GLCM. The number of misclassified mammogram was less for combined feature related classification.

Table.4: Confusion Matrix for SVM Classifier with FOS + GLCM + GLRLM Features.

Actual	predicted	
	positive	negative
Positive	true positive (tp)	false positive (fp)
25	24	1
negative	false negative (fn)	true negative (tn)
25	2	23

The suggested approach is tested by processing mammogram which is the golden standard for breast cancer screening. The mammograms are accessed from DDSM database which contains information regarding final diagnosis against each mammogram. The outcome of each testing was evaluated by means of performance measures such as sensitivity, accuracy and specificity. Fig.6 shows SVM classifier performance for the above three experiments in terms of accuracy, specificity and sensitivity. GLCM outperforms GLRLM.

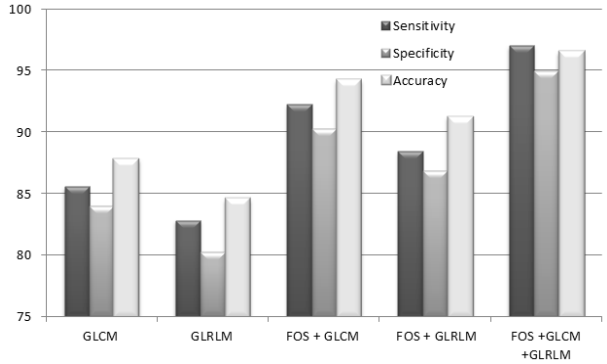


Fig. 6: SVM Classification Perfomance Comparison for Different Feature Extraction Methods.

The combination of three feature extraction method eliminates generation of false negatives and positives. The combination of first order, GLCM and GLRLM provide 95% specificity with 97% accuracy. Sensitivity for this method is 96%. The present study suggested a CAD system for automatic detection of breast cancer by classification and feature extraction of micro calcification clusters. Three feature algorithms are used such as GLCM, GLRLM and first order statistics. Contrast is foremost descriptor for all mammogram, GLCM provide contrast, energy and homogeneity values thereby it proves better results in breast cancer detection. GLRLM features are SRE, LRE and RLN. Combination of first order with GLCM is better than that of GLRLM. Accuracy of breast cancer detection can be improved by the proposed system which uses both first and second order feature set for classification. Neetha Jog et al [29] suggested similar study related to proposed work in which SVM classifier used with Gabor and GLDM feature extraction algorithms. MIAS database is used for analysis of this study. The outcomes showed 95.83% accuracy for SVM+GLDM and 71.83% accuracy for SVM + Gabor filter. These GLDM and Gabor filter can improve accuracy of cancer detection by adding them to proposed system. Further studies related to nature of breast calcification can result in design of high efficient automatic system, which even detects stages of breast cancer growth. The application of Local Binary Pattern technique can extract fine details from mammogram [30]. It is one of the main descriptor for early detection of breast cancer. From this study, it is clear that SVM is an ideal alternative as a classifier for type of cancer detection.

4. Conclusions

The proposed framework of micro calcification detection is based on SVM classifier with three feature extraction methods. The proposed system is effective in reducing misclassification rates with less training set. The limitation of system is poor performance for low contrast images and the lack of texture feature values in decision process. Future work of this paper is to implement more feature extraction methods and can use other classifiers to improve the accuracy of breast cancer detection.

References

[1] X. Zhou, R. Gordon; "Detection of early breast cancer: an overview and future prospects", *Crit. Rev. Biomed. Eng*, Vol. 17, (1989), pp. 203–255, available online: https://europepmc.org/abstract/med/2673660.

[2] S.L. Olson, B.W. Fam, P.F. Winter, F.J. Scholz, A.K. Lee, S.E. Gordon; "Breast calcifications: analysis of imaging properties", *Radiology*, Vol. 169, No.2, (1988), pp. 329–332, available online: https://pubs.rsna.org/doi/abs/10.1148/radiology.169.2.3174980.

[3] Mohapatra P J, Nanda P K, Panda S; "Color Image Segmentation Using MRF Model and Simulated Annealing", *Proceedings of Soft-computing technique for Engineering Applications*, (2006), pp. 1-10, http:// http://dspace.nitrkl.ac.in/dspace/handle/2080/357.

[4] Dana Cobzas, Neil Birkbeck, Mark Schmidt, Martin Jagersand; "3DVariational Brain Tumour Segmentation Using a High Dimensional Feature Set", *IEEE 11th International Conference on Computer Vision*, (2007), pp.1-8, http://doi.ieeecomputersociety.org/10.1109/ICCV.2007.4409130.

[5] Albert Law KW, Law FK, Francis Chan HY; "A Fast Deformable Region Model for Brain Tumor Boundary Extraction', Engineering in Medicine and Biology", *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society (EMBS/BMES) Conference*, Vol.2, (2002), pp.1055-1056, http://dx.doi.org/ 10.1109/IEMBS.2002.1106273.

[6] Azadeh Yazdan Shahmorad, Hamid Soltanianzadeh, Reza Zoroofi A; "MRSI– Braintumor characterization using Wavelet and Wavelet packets Feature spaces and Artificial Neural Networks", *Engineering in Medicine and Biology Society, 26th Annual International Conference of the IEEE*, Vol.1, No. 1-5, (2004), pp.1810-1813, http:// dx.doi.org/ 10.1109/IEMBS.2004.1403540.

[7] K.S. Woods, C.C. Doss, K.W. Bowyer, J.L. Solka, C.E. Priebe, W.P. Kegelmeyer; "Comparative evaluation of pattern recognition techniques for detection of microcalcifications in mammography", *Int. J. Pattern Recognition Artif. Intell*, Vol. 7, (1993), pp.1417–1436, available online: https:// www.worldscientific.com /doi /abs /10.1142 /S0218001493000698

[8] B. Zheng, W. Qian, L.P. Clarke; "Digital mammography: mixed feature neural network with spectral entropy decision for detection of microcalcifications", *IEEE Transactions of Medical Imaging*, Vol. 15, No.5, (1996), pp. 589–597, available online: https://ieeexplore.ieee.org /abstract /document /538936 / 10.1109/42.538936.

[9] W.J.H. Veldkamp, N. Karssemeijer; "An improved method for detection of microcalcification clusters in digital mammograms", *The SPIE Conference on Image Processing*, Vol. 3661, (1999), pp. 512–522, https://doi.org/10.1117/12.348607.

[10] S. Bothorel, B.B. Meunier, S. Muller; "A fuzzy logic based approach for semiological analysis of microcalcifications in mammographic images", *Int. J. Intelligent Systems*, Vol. 12, (1997), pp. 819–848, available online: https:// onlinelibrary.wiley.com /doi /abs /10.1002/(SICI)1098-111X(199711/12)12:11 /12%3C819 :: AID-INT3%3E3.0.CO;2-%23.

[11] Ball J.E, Bruce L.M; "Digital Mammogram Spiculated Mass Detection and Spicule Segmentation using Level Sets", *In: Proceedings of the 29th Annual International Conference of the IEEE EMBS, Cité Internationale, Lyon, France*, Vol.12, (2007), pp.4979–4984, http:// dx.doi.org/ 10.1109/IEMBS.2007.4353458.

[12] Timp S, Karssemeijer N; "Interval Change Analysis to Improve Computer Aided Detection in Mammography", *Medical Image Analysis 10*, Vol.17, (2006), pp.82–95, available online:

https://www.sciencedirect.com/science/article/pii/S1361841505000484.

[13] Fauci F, Bagnasco S, Bellotti, R., Cascio D, Cheran S.C, De Carlo, F., De Nunzio, Fantacci, M.E., Forni, G., Lauria, A; "Mammogram Segmentation by Contour Searching and Massive Lesion Classification with Neural Network", *In Proceedings: IEEE Nuclear Science Symposium Conference Record, Rome, Italy*, Vol. 5, No.16, (2004), pp. 2695–2699, http:// dx.doi.org/ 10.1109/NSSMIC.2004.1462823.

[14] Rangayyan R.M, Guliato D, Carvalho, Santiago J.D; "Feature Extraction from the Turning Angle Function for the Classification of Contours of Breast Tumors", *In: IEEE Special Topic Symposium on Information Technology in Biomedicine, Iaonnina, Greece*, Vol.8, No.23, (2006), pp. 56-59, available online: https://pdfs.semanticscholar.org/1db7/bb353c675d17a63792ff5e610826b60e1a2f.pdf.

[15] Joseph Peter V, Karnan M; "Medical Image Analysis Using Unsupervised and Supervised Classification Techniques", *International Journal of Innovative Technology and Exploring Engineering*, Vol. 3, No.5, (2013), pp. 40-45., avaiable online: https://pdfs.semanticscholar.org/1ba9/d67c80b6a762c11b9d519367e9e13a9c5c4f.pdf.

[16] Komal Sharma, AK winder Kaur, Shruti Gujral; "Brain Tumor Detection based on Machine Learning Algorithms", *International Journal of Computer Applications*,Vol. 103, No.1, (2014), pp.8875 – 8887, available online: https:// pdfs.semanticscholar.org /c102 /9db73a242801430602c108786cbe6796b7f2.pdf.

[17] Blagojce Jankulovski, Ivan Kitanovski, Katarina Trojacanec, Ivica Dimitrovski, Suzana Loskovska; "Mammography Image Classification Using Texture Features", *Proceedings of 9th Conference for Informatics and Information Technology*, Vol. 86, No.112, (2012), pp.129-132, available online: https:// pdfs.semanticscholar.org /51bc/ 6966dfc03bf18f5979acaba30adaaa61e6c2.pdf.

[18] Wu X, Kumar, Ross Quinlan J, Ghosh J, Yang Q, Motoda H, McLachlan GA, Liu P; "Top 10 algorithms in data mining", *Knowledge and Information Systems*, Vol. 14, No. 1, (2008), pp. 1–37, available online: www.cs.uvm.edu /~icdm/algorithms /10Algorithms-08.pdf

[19] Kavithal R.K, DoraiRangasamy D; "Predicting Breast Cancer Survivability Using Naïve Baysein Classifier and C4.5 Algorithm", *Elysium journal Engineering Research and management*,Vol. 25, No.42, (2014), pp.115-120, available online: *http://*ejerm.com/.../Predicting_Breast_Cancer_Survivability_Using_Naïve_Baysein_Classifier.

[20] Krishnaveni, R.Bhanumathi, T.Pugazharasan; "Study of Mammogram Microcalcification to aid tumour detection using Naive Bayes Classifier", *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* ,Vol. 3, No. 3, (2014), pp.8274-8282, available online: http:// www.siteadvisor.com /restricted.html?domain= http:%2F%2Fwww.rroij.com%2Fopen-access%2 Fstudy-of-mammogram-microcalcification-toaid-tumour-detection-using-naïve.

[21] Wei L, Yang Y, Nishikawa R.M, Jiang Y; "A Study on Several Machine-Learning Methods for Classification of Malignant and Benign Clustered Microcalcifications", *IEEE Transactions on Medical Imaging*, Vol. 24, No. 10, (2005), pp. 1278-1285, available online: https://ieeexplore.ieee.org/document/1397824/ 10.1109/TMI.2004.842457.

[22] Vijayarani, Muthulakshmi M; "Evaluating The Efficiency Of Rule Techniques For File Classification", *International Journal of Research in Engineering and Technology*, Vol. 20, No.18, (2013), pp: 2321- 2329, available online: https://pdfs.semanticscholar.org /1f33/1854e02204786f882a0492693c9b368ec9e7.pdf

[23] Dheeba J and Wiselin Jiji.G; "Detection of Microcalcification Clusters in Mammograms using Neural Network", *International Journal of Advanced Science and Technology*, Vol. 19, No.34, (2010), pp.13-22, available online: www.sersc.org/journals/IJAST/vol19/2.pdf.

[24] "USF digital mammography home page", available online: http: //marathon.csee.usf.ecu/Mammography/Database.html,(accessed May 18, 2009).

[25] Liu CC, Tsai CY, Liu J, Yu CY, Yu SS; "A pectoral muscle segmentation algorithm for digital mammograms using Otsu thresholding and multiple regression analysis", *Computer and Mathematics with Applications*, Vol. 64, No. 5, (2012), pp.100-1107, available online: https:// www.sciencedirect.com /science /article/pii/S0898122112002337.

[26] Ferrari RJ, Rangayyan RM, Desautels J E, Borges RA, Frere AF; "Automatic identification of the pectoral muscle in mammograms", *IEEE Transactions on Medical Imaging*, 2004; Vol. 23, No. 2, (2004), pp.232-245, available online: https://ieeexplore.ieee.org/document/1263612.

[27] R. C. Gonzalez, R. E. Woods, S. L. Eddins; "Digital Image Processing Using MATLAB," *Pearson Prentice Hall*, Vol. 11, (2004).

[28] J.K. Kim, H.W. Park; "Statistical textural features for detection of microcalci"cations in digitized mammograms", *IEEE Trans. Med. Imag*, Vol. 18, No. 3, (1999), pp. 231–238, available online: https://ieeexplore.ieee.org/document/764896.

[29] Neeta Jog, Arvind Pandey; "Implementation of Segmentation and Classification Techniques for Mammogram Images", *IOSR Journal of Engineering (IOSRJEN)*, Vol. 5, No. 2, (2015), available online: http://www.rroij.com /peer-reviewed/ implementation-of-segmentation-and-classification-techniques-for-mammogram-images-50187.html.

[30] S. Naresh, S. Vani Kumari; "Breast Cancer Detection using Local Binary Patterns", *International Journal Of Computer Applications*, Vol. 123, No.16, (2015), pp. 6-9, available online: https://pdfs.semanticscholar.org/137b/70428ab323da303d8d0b107c7be7571a08e3.pdf.