Comparative Analysis of Region of Interest of Different Sizes for Breast Density Classification

Vipul Sharma*

Department of Computer Science and Engineering, DAV University, Jalandhar, India
*Corresponding e-mail: vipuls85@gmail.com

ABSTRACT

In medical field, radiologists are more interested in region of interest (ROI) rather than whole image. ROI is a subpart of the image that contains very important information related to the diagnosis. In addition, ROI size has been known to influence the sensitivity and specificity of the classification. Many researchers in the past have used ROI for texture analysis of fatty and dense mammogram. Since a ROI is used as 'representative' of the image and all further computations and diagnosis depends upon the ROI, therefore, it is very crucial to select an appropriate image area as ROI. In this work, experiments have been conducted to find the appropriate size of ROI for breast density classification. Comparison of different ROI sizes 50 × 50, 75 × 75, 100 × 100, 125 × 125, 150 × 150, 175 × 175, 200 × 200, 225 × 225, 250 × 250, 275 × 275 and 300 × 300 pixels for breast density classification has been done. Different texture models have been used to extract features from these ROIs to have reliable analysis for appropriate size of the ROI. The effect of ROI size is evaluated in terms of the performance of differentiating between the fatty and dense breast tissue. For carrying out the experiments standard publicly available mini-MIAS database has been used. After the analytical study, it has been observed that a square shaped ROI having size of 200 × 200 pixels is optimal and it should be taken from the central breast region immediately behind the nipple, as this region is the densest region of the breast excluding pectoral muscle. The optimal ROI size also reduces the computational cost in extracting the texture features from the small sized ROI. The experimental results encourage the use of 200 × 200 pixels ROI for the classification of breast density.

Keywords: Region of interest, breast density, classification, texture features, feature selection

INTRODUCTION

In general population, mammographic breast density was found to be associated with an increased risk of breast cancer, as high density makes it hard for radiologists to see an abnormality which leads to false negative results [1-3]. Radiologists usually estimate breast density by visual analysis of mammogram that is considered to be highly subjective. Automatic breast density classification methods attempt to mimic such visual judgment and classify mammograms on the basis of underlying texture characteristics.

Many researchers in the recent past have used ROI for texture analysis of different abnormalities of breast [4-6]. As the subsequent calculations and analysis significantly depends upon the ROI size, thus the 'size and location' of the ROI is very important. The study of Li et al. [7], has concluded that the performance of classification significantly decreases as the ROI location was varied from the central region behind the nipple. But their study failed to show any concluding remark about the effect of ROI size on classification problem. The classification performance of breast density largely depends upon the ROI size, location and feature extraction methods adopted [8].

Finding an optimal size for breast density classification is always a challenging task. Therefore, in the present study, an investigation on ROI size has been done for the classification of breast density. A total of 45 texture features using different texture models i.e. Haralick's spatial gray level co-occurrence matrix (SGLCM) [9], gray level difference statistics (GLDS) [10], first-order statistics (FoS) [11], statistical feature matrix (SFM) [12], Law's texture energy measures (TEM) [13], fractal features [14], fourier power spectrum (FPS) [15] and are extracted from the different

sized ROIs. Feature selection is carried out by correlation based Feature Selection (CFS) [16]. For classifying the mammograms sequential minimal optimization classifier [17] is used.

During last few years, a large number of studies have been done by various researchers in order to classify mammograms on the basis of breast density [18,19]. These studies can be differentiated on the basis of difference in the area or region from which the texture features are extracted, texture features used for representing the breast density and the classifier used.

In the field of computer aided diagnosis (CAD) systems, Oliver et al. characterized breast densities using morphological and texture features in order to propose a CAD system for breast density classification [20]. A set of 322 mammographies was obtained from MIAS database and 831 mammographies were obtained from DDSM database. A decision tree classifier, k-nearest neighbour and a combined Bayesian classification were used for classification and the best results were around 82% of correct classification for the MIAS database, and 77% for DDSM database.

Vallez, et al. proposed a novel weighted voting tree classification scheme for breast density classification [21]. The authors have compared several classification methods and a novel hierarchical classification procedure of combined classifiers with linear discriminant analysis (LDA) is proposed as the best solution to classify the mammograms into the four BIRADS tissue classes. The classification scheme is based on 298 texture features. The performance is evaluated using 322 images of mini-MIAS database. The experimental results show that 99.75% of samples were correctly classified.

Li, et al. performed computerized texture analysis on mammographic parenchymal patterns for assessing breast cancer risk [7]. Seventeen features, which characterize the density and texture of the parenchymal patterns, were extracted from the ROIs on these digitized mammograms. Stepwise feature selection and linear discriminant analysis were applied to identify features that differentiate between the low-risk women and the BRCA1/BRCA2 gene-mutation carriers. Receiver operating characteristic (ROC) analysis was used to assess the performance of the features in the task of distinguishing between these two groups. A total of 90 mammograms taken from cancer risk clinics at University of Chicago and University of Pennsylvania were used for carrying out the work. Their results show that there was a statistically significant decrease in the performance of the computerized texture features, as the ROI location was varied from the central region behind the nipple.

Bovis, et al. proposed an approach for the classification of mammographies on the basis of breast density [22]. The authors investigated the use of four texture models i.e. spatial grey level dependency matrices (SGLD), fourier power spectrum, Law's texture measures, discrete wavelet transform for classifying mammograms. They evaluated two different tasks: four-class classification problem following the breast imaging-reporting and data system (BI-RADS) standard and two-class problem that classify mammogram into fatty or dense types. A total of 377 mammograms from digital database for screening mammography (DDSM) are selected to evaluate the performance.

Oliver, et al. proposed a CAD system for classification of breast density using morphological and texture features [20]. A set of 322 images from mini-MIAS database and 831 images from DDSM database were used for evaluating the performance of the proposed system. For classification, a decision tree classifier, Bayesian classifier and K-nearest neighbour were used.

Petroudi, et al. presented an approach for automatic classification of breast density patterns [23]. These breast density patterns within the breast are modelled by statistical features. In their method, each mammogram is segmented into three different components: breast tissue, background and pectoral muscle. Finally, a total of 132 mammograms selected from the Oxford database are classified into BI-RADS standard.

In most of the previous studies, for the classification of breast density the entire breast was processed to extract the texture features.

Rouhi, et al. proposed two automated methods are presented to diagnose mass types of benign and malignant in mammograms [24]. In the first proposed method, segmentation is done using an automated region growing whose threshold is obtained by a trained artificial neural network (ANN). In the second proposed method, segmentation is performed by a cellular neural network (CNN) whose parameters are determined by a genetic algorithm (GA). Intensity, textural, and shape features are extracted from segmented tumours. GA is used to select appropriate features from the set of extracted features. In the next stage, ANNs are used to classify the mammograms as benign or malignant. For evaluating the performance, the

authors have used 93 mammograms from MIAS and 179 mammograms from DDSM database.

Ergin, et al. proposed a framework for the classification of normal and abnormal (cancerous) breast tissues [25]. The histogram of oriented gradients (HOG), dense scale invariant feature transform (DSIFT), and local configuration pattern (LCP) methods were used to extract the rotation- and scale invariant features for all tissue patches. A classification is made by utilizing support vector machine, k-nearest neighbourhood, decision tree, and fisher linear discriminant analysis (FLDA). For experiments, 400 patches from IRMA version of DDSM database have been used by the authors.

The remaining paper is organized as follows: Section II describes the material and methods. In section III, results and discussions are presented. Finally, conclusions are given in section IV.

MATERIAL AND METHODS

Image Dataset Used

The mammograms database used to carry out this study is mini-MIAS (mammographic image analysis society) [26]. The database contains 322 images in mediolateral-oblique (MLO) view. Each image in MIAS is of size 1024×1024 pixels. The images in this database are classified in three categories based on their density as fatty, fatty-glandular and dense-glandular.

Table 1 Division of breast density classes in mini-MIAS database

Class	Character of background tissue	No. of Images
I	Fatty (F)	106
II	Fatty-glandular (G)	104
III	Dense-glandular (D)	112
Total Images		322

Table 1 shows the distribution of images according to breast density in mini-MIAS database. In this study, all the Fatty glandular and dense glandular mammograms are treated as one group of dense mammograms giving a two-class classification problem (fatty, dense). Combining images of fatty glandular into dense glandular resulted in 216 images of dense category. For evaluating the performance, 3-fold cross validation has been used five times for evaluation. In each iteration, a certain proportion i.e. two-thirds-of the data is randomly selected for training, and the remainder one-third is used for testing. The accuracy on the different iterations is averaged to yield an overall accuracy [20].

Extraction of Region of Interest

In computer aided diagnostic system, when the region of interests (ROIs) are used to extract the features then accuracy of the system significantly depends upon the characteristics of ROI. Since an ROI is used as 'representative' of the image, and all further computations and diagnosis depends upon the ROI, therefore, it is very crucial to select an appropriate image area as ROI.

The study of Li, et al. [7], has concluded that the performance of classification significantly decreases as the ROI location was varied from the central region behind the nipple. But their study failed to show any concluding remark about the effect of ROI size on classification problem. Figure 1 shows the schematic illustration of ROIs extracted from different location within breast by Li, et al. These included the central breast region immediately behind the nipple (position A), the centre of the central breast region (position D), the central left breast region far away from nipple (position E), the upper central breast region (position B) and the lower central breast region (position C). The classification performance of breast density significantly depends upon the ROI size, location and feature extraction methods adopted.
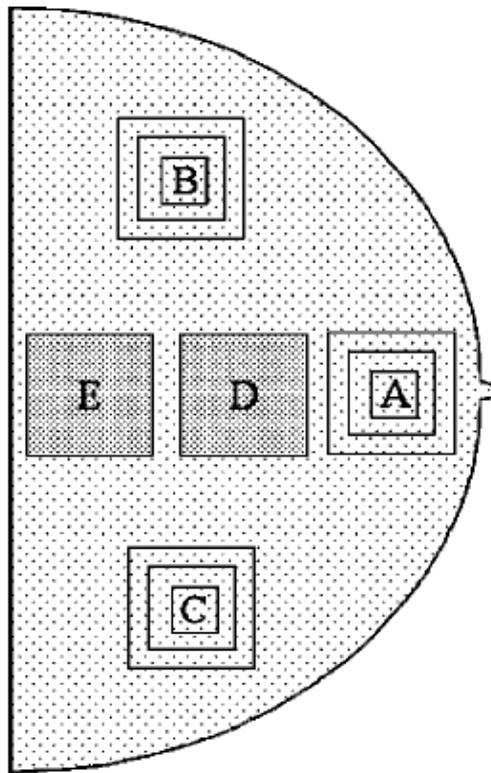
Figure 1 Schematic illustration of ROIs extracted from different location within breast by Li, et al. [7]

For carrying out the experiments, ROIs, 300 × 300 pixels in size were manually selected from the breast from position 'A' as shown in Figure 1. This included the central breast region immediately behind the nipple since that region contains the densest region of the breast as shown in Figure 2. ROIs of smaller sizes 50 × 50, 75 × 75, 100 × 100, 125 × 125, 150 × 150, 175 × 175, 200 × 200, 225 × 225, 250 × 250 and 275 × 275 pixels are extracted from the centre of each 300 × 300 pixel size ROI as shown in Figure 3.
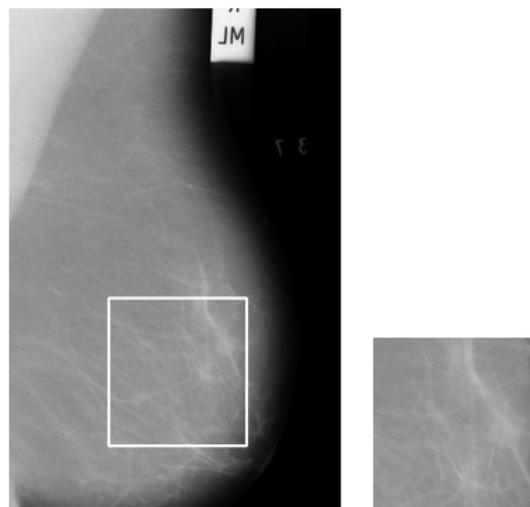
Original Image mdb006   Extracted ROI

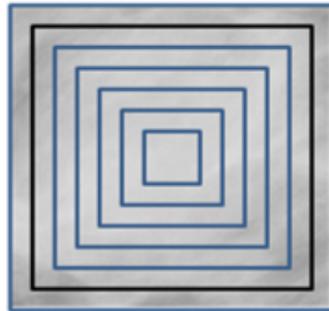**Figure 2 300 × 300 pixels ROI selected from central breast region**

Figure 3 Extraction of Sub ROIs from 300 × 300 pixels ROI

The main reason for selecting square shape ROI is that, the most of the texture models are based on matrices calculations, which could be easily done for square matrices. Thus, the square shaped ROIs have been selected for the present study.

Sample ROI's of size 200 × 200 pixels extracted from mammograms of different breast densities taken from mini-MIAS dataset are shown in Figure 4.
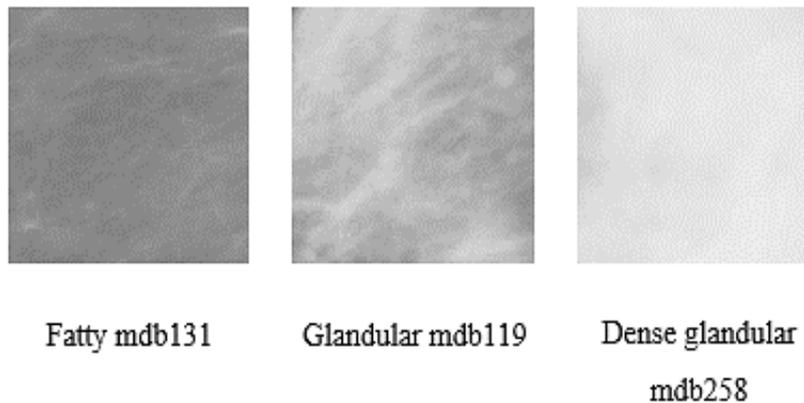

Figure 4 Sample ROIs from dataset

Texture Feature Extraction

The gray level intensities of fatty and dense breast tissues are different in nature. Various models to extract texture feature from an image exist in literature and visualize texture in a different way. Table 2 shows the features used from different texture models for classification purpose. Brief explanation of the texture models used is given below:

Table 2 Features used from different texture models

Model	Features Extracted
SGLCM (F_{1-14})	Angular Second Moment, Contrast, Correlation, Sum of Squares, Inverse Difference Moment, Sum Average, Sum Variance, Sum Entropy, Entropy, Difference Variance, Difference Entropy, Information Measure of Correlation 1, Information Measure of Correlation 2, Maximal Correlation Coefficient
GLDS (F_{15-19})	Homogeneity, Contrast, Mean, Energy, Entropy
FoS (F_{20-23})	Mean, Variance, Skewness, Kurtosis
SFM (F_{24-27})	Coarseness, Contrast, Periodicity, Roughness
Law's TEM (F_{28-41})	Edge_Level, Spot_Level, Wave_Level, Ripple_Level, Spot_Edge, Wave_Edge, Ripple_Edge, Wave_Spot, Ripple_Spot, Ripple_Wave, Edge_Edge, Spot_Spot, Wave_Wave, Ripple_Ripple
Fractal (F_{42-43})	Hurst Coefficient at Resolution 1, Hurst Coefficient at Resolution 2
FPS (F_{44-45})	Radial Sum, Angular Sum

SGLCM Features

The SGLCM based features, as proposed by Haralick, et al. are the most frequently used texture features [9]. The SGLCM is a matrix, where the number of rows and columns is equal to the number of quantized gray levels, Ng, in

the image. Co-occurrence matrices describe how often one gray level will appear in a specified spatial relationship to another gray-level in the image. Spatial relationship is a function of angular relationship θ and the distance (d) between the neighbouring pixels (resolution cells).

GLDS Features

The GLDS algorithm uses first-order statistics of local property values based on the absolute differences between pairs of gray levels or of average gray levels to extract the following 5 texture measures: Homogeneity, Contrast, Mean, Energy and Entropy [10]. These features are based on the absolute difference between pairs of gray levels separated at distance δ=(Δx, Δy).

FoS Features

In this model, the first order statistics describing the image gray-level distribution (histogram) is used. These features are derived from normalized histogram of image. Central moments derived from the histogram are used to characterize texture of the image. The following statistical features are computed on the basis of pixel's gray value in the image: mean value, variance, skewness and kurtosis [11].

SFM Features

The SFM measures the statistical properties of pixel pairs at several distances within an image, which are used for statistical analysis. Based on the SFM, the texture features like coarseness, contrast, periodicity, and roughness are computed [12].

Law's TEM Features

Another texture quantification approach that uses convolution with specialized filters is based on Laws masks [13]. These are constructed using three foundational vectors, which correspond to center-weighted Local averaging, Edge detection with symmetric first differencing, and Spot detection with second differencing, in one dimension. These measures are computed by first applying small convolution kernels to a digital image and then performing a nonlinear windowing operation.

Fractal Features

The Hurst coefficients (H(k)) are computed for different image resolutions, where a smooth texture surface is described by a large value of the parameter H(k), whereas the lower value indicates a rough texture surface [14]. In this model, Hurst coefficient is calculated at two image resolutions.

FPS Features

This texture model contains the information on the texture orientation, grain size, and texture contrast of the image. The discrete fourier transform (DFT) approach is used here for texture quantification because repetitive global patterns are difficult to describe with spatial techniques but relatively easy to represent with peaks in the spectrum [15]. The radial sum and the angular sum of the DFT were computed to describe texture.

Since texture provides information regarding the spatial distribution of gray levels along with variations in brightness [27], it can be used for breast density assessment. Texture features are extracted from each ROI selected from the mammogram. All the extracted features are normalized so as to have unit variance and zero mean.

Feature Selection

Of the forty-five texture features generated by different texture models used in this study, some features may turn out to be less informative and redundant. So, in order to select highly discriminating features, correlation-based feature selection (CFS) [16] is used. Being a simple filter algorithm CFS makes use of a correlation based heuristic function for evaluating the importance of features. It evaluates the importance of subset of features by taking into consideration two factors, namely the individual predictive ability of each feature and the degree of redundancy associated with them. Those subset of features that show high correlation with the class and also have lower inter-correlation are preferred. In this study, feature selection is done by Waikato environment for knowledge analysis (WEKA) [28] tool and search space is explored with best first search strategy using forward selection with a stopping criterion of five consecutive fully expanded non-improving subsets. Initial search starts with empty set of features that has merit 0.0.

Merit of each subset is evaluated by the heuristic function. From all the subsets evaluated, the subset with the highest merit is reported. The selected features by CFS method are presented in Table 3. The selected twelve features out of forty-five texture features are fed to the classifier as input.

Table 3 Selected features from different texture models

Model	Feature Selected
SGLCM	Correlation (F_3), Inverse Difference Moment (F_5)
GLDS	Entropy (F_{19})
FoS	Mean (F_{20}), Skewness (F_{22})
SFM	(No feature selected)
Law's TEM	Ripple Edge (F_{34})
Fractal	Hurst Coefficient at Resolution 2 (F_{43})
FPS	Angular Sum (F_{45})

Classification

The aim of classification is to classify objects into predefined categories or classes. When performing analysis on mammograms for classification of breast density, the main intention of classifier is to use the significant features obtained from the feature extraction step as input to classify the ROIs into fatty and dense categories. This automated classification is developed to further improve the decision-making capability of radiologist in making final judgment. From the literature survey, it has been found that SVM classifier has been able to attain substantially higher accuracy than traditional classification techniques. It has played an important role in medical imaging to classify mammograms with accuracy levels up to 90% [29,30]. Machine learning methods play a significant role at this stage to perform the classification and decision process [30,31]. SVM being generalized classifier is helpful in efficient classification especially in the cases where numbers of training images are quite less. In the present work, a new SVM learning algorithm called sequential minimal optimization has been used. SMO is conceptually simple, easy to implement, is generally faster, and has better scaling properties for difficult SVM problems than the standard SVM training algorithms [17]. For the classification problem, an optimal parameter setting for the classifier has been done through experimentation. In this study, the best results are obtained by polynomial kernel with degree one.

RESULTS AND DISCUSSION

The effect of ROI size is evaluated in terms of the performance of differentiating between the fatty and dense breast tissue. The performance is evaluated in terms of sensitivity, specificity, accuracy, and ROC Area. For evaluating the performance three-fold cross validation has been used. The results of classification performance with different ROI sizes are given in Table 4.

A high of these values is obtained in case of 200 × 200 pixels size ROI. From the result analysis, it has been found that as the ROI size increases from 50 × 50 pixels, there is increase in the overall classification accuracy up to size 200 × 200 pixels. The reason may be due to the fact that as the ROI size increases a sufficient portion is available for the texture analysis of fatty and dense tissue. ROIs of size 125 × 125 pixels and 150 × 150 pixels result in same overall accuracy of 93.5% but a high recognition rate is obtained for dense tissue in case of 125 × 125 pixels ROI. An overall accuracy of 95.4% is obtained in case of 175 × 175 and 250 × 250 pixels ROI. The poor performance of small sized ROIs may be due to the fact that these sizes cover only a small portion of breast tissue that is not sufficient for the analysis purpose. ROI of size more than 250 × 250 pixel results in lesser performance because in some mammograms this size covers pectoral muscle and background. Table 4, shows that a high ROC-AUC value is obtained in case of 200 × 200 pixel size ROI.

Table 4 Classification performance of different ROI sizes

ROI Size (Pixels)	Sensitivity (%)	Specificity (%)	Accuracy (%)	ROC-AUC
50 × 50	93.5	78.1	88.9	0.858
75 × 75	94.8	78.1	89.9	0.948
100 × 100	96.1	84.4	92.6	0.961
125 × 125	96.1	87.5	93.5	0.968

150 × 150	93.5	93.8	93.5	0.964
175 × 175	96.1	93.8	95.4	0.972
200 × 200	100	88.46	96.46	0.996
225 × 225	100	87.5	96.3	0.981
250 × 250	98.7	87.5	95.4	0.967
275 × 275	94.8	84.4	91.7	0.958
300 × 300	90.9	81.2	88	0.848

From the analysis of results, it has been found that, the performance of texture features varies as the ROI size is varied. The experimental results show that use of ROI of size 200 × 200 pixel results in good classification accuracy. A sensitivity of 100% is achieved in this case means all the dense cases are being classified as dense that is highly desirable. Moreover, this size covers most of the densest region of the breast excluding pectoral muscle and background for analysis. This also reduces the computational cost in extracting the texture features from the small sized ROI.

However, ROIs of size less than 200 × 200 pixels are smaller in size, cover small portion of the dense tissue and do not have sufficient information to get a reliable statistical analysis from them. These small sized ROIs may be useful in case of cyst or small size tumour in breast but not suitable for the present study. It has been observed that ROI of size more than 250 × 250 pixels is not suitable for present study as in some cases this size covers some region of pectoral muscle and background that is not suitable for analysis purpose.

CONCLUSION

The accuracy of CAD system significantly depends upon the characteristics of the ROI. Tissue characteristics vary according to the location and size of ROI, which has a considerable impact on the classification performance. Therefore, in the presented study, experiments were conducted to find the optimal ROI size for breast density classification. For carrying out the experiments, a wide range of texture models have been used and feature selection has been done to select the highly informative features from the extracted features. The experiments were conducted on mini-MIAS dataset.

After the analytical study, it has been observed that a square shaped ROI having size of 200 × 200 pixels is optimal for mini-MIAS database and it should be taken from the central breast region immediately behind the nipple, as this region is the densest region of the breast excluding pectoral muscle. The optimal ROI size also reduces the computational cost in extracting the texture features from the small sized ROI.

The results prove that the use of various texture models, feature selection and ROI of size 200 × 200 pixels has significantly improved the classification performance. The results encourage the use of proposed set of optimal features and 200 × 200 pixels ROI for classifying mammogram in one of the two (Fatty, Dense) breast density classes. In the present work, more importance is given for the classification of two categories namely Fatty and Dense. This work can be extended for the classification of breast density on the basis of BI-RADS standard in near future.

REFERENCES

[1] Wolfe, John N. "Breast patterns as an index of risk for developing breast cancer." *American Journal of Roentgenology* 126.6 (1976): 1130-1137.

[2] Boyd, Norman F., et al. "Mammographic breast density as an intermediate phenotype for breast cancer." *The lancet oncology* 6.10 (2005): 798-808.

[3] Aiello, Erin J., et al. "Association between mammographic breast density and breast cancer tumor characteristics." *Cancer Epidemiology and Prevention Biomarkers* 14.3 (2005): 662-668.

[4] de Oliveira, Julia E. E., et al. "MammoSys: A content-based image retrieval system using breast density patterns." *Computer methods and programs in biomedicine* 99.3 (2010): 289-297.

[5] Deserno, Thomas M., et al. "Computer-aided diagnostics of screening mammography using content-based image retrieval." *SPIE medical imaging*. International Society for Optics and Photonics, 2012.

[6] de Oliveira, Júlia Epischina Engrácia, Arnaldo de Albuquerque Araújo, and Thomas M. Deserno. "Content-based image retrieval applied to BI-RADS tissue classification in screening mammography." *World journal of radiology* 3.1 (2011): 24.

[7] Li, Hui, et al. "Computerized analysis of mammographic parenchymal patterns for assessing breast cancer risk: Effect of ROI size and location." *Medical Physics* 31.3 (2004): 549-555.

[8] Huo, Zhimin, et al. "Computerized analysis of mammographic parenchymal patterns for breast cancer risk assessment: Feature selection." *Medical Physics* 27.1 (2000): 4-12.

[9] Haralick, Robert M., and Karthikeyan Shanmugam. "Textural features for image classification." *IEEE Transactions on systems, man, and cybernetics* 3.6 (1973): 610-621.

[10] Weszka, Joan S., Charles R. Dyer, and Azriel Rosenfeld. "A comparative study of texture measures for terrain classification." *IEEE Transactions on Systems, Man, and Cybernetics* 4 (1976): 269-285.

[11] Bankman, Isaac, editor. *Handbook of medical image processing and analysis*. academic press, 2008.

[12] Wu, Chung-Ming, and Yung-Chang Chen. "Statistical feature matrix for texture analysis." *CVGIP: Graphical Models and Image Processing* 54.5 (1992): 407-419.

[13] Laws, Kenneth I. "Rapid texture identification." *24th annual technical symposium*. International Society for Optics and Photonics, 1980.

[14] Wu, Chung-Ming, Yung-Chang Chen, and Kai-Sheng Hsieh. "Texture features for classification of ultrasonic liver images." *IEEE Transactions on medical imaging* 11.2 (1992): 141-152.

[15] Lendaris, Geoge G., and Gordon L. Stanley. "Diffraction-pattern sampling for automatic pattern recognition." *Proceedings of the IEEE* 58.2 (1970): 198-216.

[16] Hall, Mark A., and Lloyd A. Smith. "Feature selection for machine learning: Comparing a correlation-based filter approach to the wrapper." *FLAIRS conference*. Vol. 1999. 1999.

[17] Platt, John. "Sequential minimal optimization: A fast algorithm for training support vector machines." (1998).

[18] Sharma, Vipul, and Sukhwinder Singh. "CFS-SMO based classification of breast density using multiple texture models." *Medical and biological engineering and computing* 52.6 (2014): 521-529.

[19] Sharma, Vipul, and Sukhwinder Singh. "Automated classification of fatty and dense mammograms." *Journal of Medical Imaging and Health Informatics* 5.3 (2015): 520-526.

[20] Oliver, Arnau, et al. "A novel breast tissue density classification methodology." *IEEE Transactions on Information Technology in Biomedicine* 12.1 (2008): 55-65.

[21] Vállez, Noelia, et al. "Breast density classification to reduce false positives in CADe systems." *Computer methods and programs in biomedicine* 113.2 (2014): 569-584.

[22] Bovis, Keir, and Sameer Singh. "Classification of mammographic breast density using a combined classifier paradigm." *4th international workshop on digital mammography*. 2002.

[23] Petroudi, Styliani, Timor Kadir, and Michael Brady. "Automatic classification of mammographic parenchymal patterns: A statistical approach." *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*. Vol. 1. IEEE, 2003.

[24] Rouhi, Rahimeh, et al. "Benign and malignant breast tumors classification based on region growing and CNN segmentation." *Expert Systems with Applications* 42.3 (2015): 990-1002.

[25] Ergin, Semih, and Onur Kilinc. "A new feature extraction framework based on wavelets for breast cancer diagnosis." *Computers in biology and medicine* 51 (2014): 171-182.

[26] Suckling, John, et al. "The mammographic image analysis society digital mammogram database." *Exerpta Medica. International Congress Series*. Vol. 1069. 1994.

[27] Gonzalez, Rafael C., and Richard E. Woods. "Image processing." *Digital image processing* 2 (2007).

[28] Hall, Mark, et al. "The WEKA data mining software: An update." *ACM SIGKDD explorations newsletter* 11.1 (2009): 10-18.

[29] El-Naqa, Issam, et al. "Content-based image retrieval for digital mammography." *Image Processing. 2002. Proceedings. 2002 International Conference on*. Vol. 3. IEEE, 2002.

[30] Pujari, Devashish, Rajesh Yakkundimath, and Abdulmunaf S. Byadgi. "SVM and ANN based classification of plant diseases using feature reduction technique." *IJIMAI* 3.7 (2016): 6-14.

[31] Esmaeilpour, Mansour, and Ali Reis Ali Mohammadi. "Analyzing the EEG signals in order to estimate the depth of anesthesia using wavelet and fuzzy neural networks." *International Journal of Interactive Multimedia and Artificial Inteligence* 4.Regular Issue (2016).