Characterization of Normal Breast Tissues Using Texture Analysis

M. E. M. Gar elnabi\textsuperscript{1}, M. J. Alport\textsuperscript{2}, W. I. D. Rae\textsuperscript{3}, M. A. A Omer\textsuperscript{1}, and E. A. A. Balla\textsuperscript{1}

\textsuperscript{1}College of Medical Radiologic Science, Sudan University of Science and Technology, Khartoum, Sudan
\textsuperscript{2}School of Physics, University of KwaZulu-Natal, Durban, South Africa.
\textsuperscript{3}Department of Medical Physics, University of Free State, Bloemfontein, South Africa.

**ABSTRACT:** The identification of normal breast tissues in mammograms is an important step in identifying abnormal tissues, or masses. The major focus of this study is to use image-processing techniques to characterize the normal breast tissues. Clustering techniques were developed to classify the breast tissue into one of 5 different tissue types including fat, glandular, connective, dense tissues and pectoral muscle. The classification was achieved by using a novel technique involving a subset of the elements (the values on the diagonals) in the Spatial Grey Level Dependence (SGLD) matrix which was calculated using texture features, thus it was called the diagonal SGLD method (dSGLD). The ultimate goal of this study was therefore to design and implement a computerized characterization method that mimics the radiologist’s characterization of tissue composition in digitized mammograms. The classification results using dSGLD and SGLD features achieved an accuracy of 80\% and 71\%, respectively when compared to radiologists’ classifications.

**KEYWORDS:** SGLD, Characterization of normal breast tissue, Mammogram

**INTRODUCTION**

Although the human breast can be imaged with high spatial resolution using X-rays, mammograms are often difficult to interpret. The female breast has a complex architecture containing many types of soft tissue. The structure of the breast varies during the lifetime of a woman with development at puberty, cyclic modifications during the period of sexual activity and postmenopausal involution. As a consequence the borders between regions of the breast in mammograms are difficult to identify unambiguously and change at different times during the patient’s lifetime\textsuperscript{(1,2)}. Densities of some mammographic structures are low and there is an overlap between different normal and pathological tissues. Due to the nature of planar radiology each pixel of a digitized mammogram represents several overlying tissues; with the net pixel intensity being determined by the X-ray attenuation through a mixture of the underlying different tissue types. Often one tissue type dominates attenuation within a pixel and it can be said to represent that particular tissue. The normal female breast contains a number of tissues which contribute different intensities in mammographic images due to varying amounts of fatty, glandular, and connective components. Usually there is a fat layer just beneath the skin, below which lies glandular tissue supported by a network of fibrous connective tissue. This set of tissue constituents means that composition of the breast is extremely variable. Radiolucent adipose tissues form a large portion of the breast\textsuperscript{(3,4)}. Other tissues such as ducts, lobular elements and fibrous connective tissues are also visible due to their different radiographic densities. Ducts are typically seen as thin linear structures radiating back from the nipple and these criss-crosses...
other tissues. Glandular lobules and intra-lobular connective tissues appear as vague fluffy densities within the mammogram. Connective tissues are of two types: interlobular and extra-lobular. Interlobular connective tissues aid the visibility of the lobule. The majority of the radiographic density on a mammogram can be attributed to the extra-lobular connective tissues. Although classification and characterization of breast tissues in a mammogram is generally thought to be useful for breast imaging reports, the task remains both subjective and relatively crude because of the large inter- and intra-observer variability. To overcome some of these limitations, a computerized approach was implemented in this study, which provides a quantitative, objective, and reproducible assessment of normal breast tissue composition. Texture features have been used in several studies to characterize the tissues in digitized mammograms as normal or malignant. Other studies have focused on characterizing the normal breast tissues with a varying degree of success. The first and second-order grey level histograms have been used to classify the mammogram into fibro-glandular fat tissues. These results showed that the correlation coefficient between the area and perimeter of the region that was outlined by the radiologist, versus the algorithm, was 0.74 and 0.59 respectively. Granulometric and Laws’ texture masks have been used to classify the mammograms into fatty and glandular regions. A recognition rate of 75.7% for granulometric and 80.3% for Laws’ masks were reported. An optimal threshold seeking algorithm was used to segment and quantify the dense regions in the breast using a variation image. These results had a correlation coefficient ranging from 0.92 to 0.95 when compared with the mammographer’s estimation. The Fourier transform, spatial relationship and features based on the absolute values of grey level have been used, to classify the whole mammogram into one of the 4 BI-RADS categories with a classification accuracy of 66.8%. Similarly achieved a classification accuracy of 67%, while Chang et al. used features extracted from image statistics, variation, mathematical morphology, textural edginess, and Gaussian subtraction, to generate a computer index that mimicked the radiologist rating of tissue composition with a correlation coefficient of 0.87 when compared to the radiologist scoring.

MATERIALS and METHODS
The mammograms used in this study were randomly selected from the patient files in the Department of Radiology, Addington Hospital, Durban, South Africa. Sixty one normal mammograms were selected. Nine mammograms were used as the training set and 52 mammograms as the test set, to investigate the classification accuracy of texture features extracted from the Spatial Grey Level Dependence (SGLD) matrix, as well as to evaluate the classification accuracy of a new technique using the diagonal elements of the Spatial Grey Level Dependence (dSGLD) matrix. The classification was carried out assuming five tissue classes: fat, glandular, connective, dense and pectoral muscle. The mammograms were digitized with a laser film scanner (Epson Expression 1640XL, model EU-22, # 015119, Seiko Epson Corporation, Japan) at 600 dpi. The digitized mammograms were saved in image
files using the lossless TIFF format. The data was analyzed by a PC Pentium IV 1.28 GHz processor using IDL1.

**Scoring of breast tissue composition by the Radiologists:** The Four radiologists that participated in this study were asked to classify regions in the breast tissues in the mammograms as predominantly fatty, glandular, fibrous or dense tissues. They drew borders around the selected regions on a transparency that was overlaid on the mammogram. The transparency was marked with labels and corners, so that these could be aligned with the mammogram labels and the corners of the x-ray film for registration purposes.

**Spatial Grey level Dependence (SGLD) matrix:** Fourteen textural features were examined, including entropy, energy, inertia, inverse difference moment, sum entropy, difference entropy, correlation, sum average, sum variance, difference variance, difference average, information measure of correlation1, information measure of correlation2 and variance. The textural features were computed from the SGLD matrices that were constructed in four directions (θ = 0°, 45°, 90°, and 135°) and a distance, \( d \), of 1 pixel, giving a total number of 56 features (14 features × 4 angles). The distance, \( d = 1 \), was used to ensure that a large number of occurrences were derived from the underlying structures. Features computed from the SGLD matrices with larger distances (\( d = 3, 5 \) and 10) have been found \(^{16}\) to possess a higher degree of correlation (≥0.9) than corresponding features for \( d = 1 \). For this reason, in this study, breast tissue classifications were carried out using features that were extracted from SGLD matrices constructed with a distance, \( d = 1 \) pixel, in order to reduce correlation between the calculated texture features.

Two approaches were investigated to select the optimum subset of SGLD textural features and are presented here. The first approach involved an exhaustive search for the optimal subset of texture features based on the classification accuracy of single and combined features. The second approach used stepwise linear discriminant analysis to select the optimal subset of texture features.

**RESULTS**

\( F_{50} \) features: In the exhaustive search approach, the texture features were averaged over the four directions; thus reducing the number of features from 56 to 14. The mammograms were then classified using single texture feature classification. Texture features which produced a classification accuracy of more than 50% were then combined in all possible ways and the classifications were repeated for the same mammograms using this set, which were then referred to as the set of \( F_{50} \) features, which included entropy, inverse difference moment, sum average, sum entropy, difference entropy and difference average. The classification accuracy was evaluated for each classification map by calculating a confusion matrix using the radiologists’ classification as ground truth. The combination of all the \( F_{50} \) features gave the best results.

SLD analysis: Stepwise linear discriminant analysis (SLD), was used to investigate the impact of directionality on the classification accuracy, by selecting an optimal subset of the 56 texture features that were computed from the training set, using the statistical package SPSS for Windows Version 11.0. The selected features (\( F_{SLD} \)) were then used to

---

1 Interactive Data Language Version 5.6.
classify the training set into 5 classes using the K-means algorithm. The class centres were then used to classify the unseen mammograms assuming 5 tissue classes based on the minimum distance from the class centres.

dSGLD features: The sum of the probabilities in the diagonal entries of the SGLD matrix referred to as the dSGLD elements can be computed as follows:

\[ F_l = \sum_{i=1}^{M} \sum_{j=l}^{M} P_{d,l} (i, j) \]  

where \( F_l \) represents the sum of diagonal probabilities on diagonal, \( l \), for \( l = 0, 1, 2, \ldots, M \), where \( M \) is the \( M \)th diagonal.

The most discriminating diagonals, using stepwise linear discriminant analysis, were the main and adjacent diagonals, i.e. (\( D_0 \) and \( D_1 \)). Figure 1 shows the probability distributions of the first 10 diagonals of the 5 tissue types. The values represent the average probabilities on each diagonal using 3312 categorized sub-images. The mean probability and their corresponding standard deviations for each of those diagonals were obtained for each class for classification purposes as shown in Table 1.

In order to classify the normal breast tissues in the digitized mammograms using, \( F_l \), of dSGLD, into 5 classes, the normalized Euclidian distance, \( N_d \), was applied. The classification was done according to the minimum distance from the class centre, which was obtained from the training set using the mean values of \( F_l \) for each class. The classification accuracy using \( N_d \) was compared to the classification using Euclidian distance (\( E_d \)), Mahalanobis distances (\( M_d \)) and linear discriminant analysis (\( L_d \)).

\( N_d \) was used in an attempt to enhance the performance of the classifier, since it is able to take into account the variations within the class by including the standard deviation of the respective class, as shown in Eq. 2. \( N_d \) is derived from the known \( E_d \). To compute the \( N_d \), it was assumed that there were \( K \) classes of tissues, and that these classes were represented by the class centres \( z_1, z_2, \ldots, z_K \). The distance, \( N_d \), between the \( l \)th feature vector, \( F_l \), and the \( i \)th class centre, \( z_i \), is given by:

\[ N_{d,i} = \frac{1}{M} \left( \sum_{l=1}^{M} \frac{(F_l - z_{il})^2}{\sigma_{il}} \right) \]  

where: \( F_l \) is the feature vector that is calculated as the sum of the probabilities on diagonal, \( l \), \( i = 1, 2, \ldots, M \), with \( M \) being the number of diagonals,

\( z_i \) = the centre of class \( i \), where \( i = 1, 2, \ldots, K \), with \( K \) being the number of classes, and \( \sigma_i \) is the standard deviation of the \( i \)th class, for the \( l \)th diagonal.

The classifier, \( N_d \), computes the distance between the feature, \( F_l \) of the unknown class and the class centre of from 3312 sub-images, each of size 80 \times 80 pixels.
each class, and assigns the feature to the closest class if the distance is less than 1σ from the centre, else it is deemed to belong to an unknown class.

**Classification results:** To classify the unseen mammograms, the dSGLD and SGLD features were first computed from the SGLD matrices for all mammograms using a window of size 80×80 pixels. The extracting window was incrementally moved in 20 pixel steps in the x and y directions and at each step the features were calculated. In this way, these features were calculated for each overlapping 80×80 sub-image for the whole mammogram. In order to minimize the misclassification of normal breast tissue as pectoral muscle, the classification first segmented the pectoral muscle then, in the second step, the breast tissue was classified into its different types. The mammogram tissues were first classified into 3 classes (fat, fibro-glandular and pectoral muscle). This classification was based on the minimum distance from the class centres, using the corresponding class obtained from the training set using the $F_{50}$ and dSGLD features respectively. A classification map was then generated for each mammogram (Figure 2), with class values of 1, 2 and 3 assigned to fat, fibro-glandular and pectoral muscle tissues respectively. The pectoral muscle was generally located in the upper 3rd of the mammogram; therefore a predefined window located in this region was used to search for the class that represented the pectoral muscle. If pectoral muscle was found in the predefined window, then the classification map was thresholded using the class value of pectoral muscle. The threshold image generally contained the pectoral muscle and some island regions that were classified as the pectoral muscle class and were labelled using the region-labelling algorithm in IDL. The labelled region that corresponded to the pectoral muscle was selected and the other island regions were discarded. After segmenting the pectoral muscle, the remaining regions of the mammogram were classified into 4 classes; fat, glandular, connective and dense tissues, using the class centres of the dSGLD and $F_{50}$ features with the corresponding features concurrently. The feature vectors that corresponded to the pectoral muscle (if present) were extracted from the features and assigned to pectoral muscle before classifying the features that represented the breast tissues.

### Table 1: Normalised confusion matrix using (a) dSGLD and (b) SGLD features.

<table>
<thead>
<tr>
<th>Tissues</th>
<th>Fat</th>
<th>Glandular</th>
<th>Connective</th>
<th>Dense</th>
<th>Unknown</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fat</strong></td>
<td>93.4</td>
<td>0.5</td>
<td>5.0</td>
<td>0.0</td>
<td>1.2</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Glandular</strong></td>
<td>3.0</td>
<td>77.0</td>
<td>10.1</td>
<td>9.9</td>
<td>0.1</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Connective</strong></td>
<td>30.3</td>
<td>12.7</td>
<td>56.7</td>
<td>0.2</td>
<td>0.0</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Dense</strong></td>
<td>0.7</td>
<td>10.4</td>
<td>0.7</td>
<td>84.0</td>
<td>4.1</td>
<td>100.0</td>
</tr>
</tbody>
</table>

(a)

<table>
<thead>
<tr>
<th>Tissues</th>
<th>Fat</th>
<th>Glandular</th>
<th>Connective</th>
<th>Dense</th>
<th>Unknown</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fat</strong></td>
<td>86.4</td>
<td>3.7</td>
<td>9.8</td>
<td>0.1</td>
<td>0.0</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Glandular</strong></td>
<td>2.8</td>
<td>67.3</td>
<td>18.3</td>
<td>11.5</td>
<td>0.0</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Connective</strong></td>
<td>45.4</td>
<td>8.0</td>
<td>46.3</td>
<td>0.3</td>
<td>0.0</td>
<td>100.0</td>
</tr>
<tr>
<td><strong>Dense</strong></td>
<td>0.7</td>
<td>30.3</td>
<td>1.4</td>
<td>66.2</td>
<td>1.4</td>
<td>100.0</td>
</tr>
</tbody>
</table>

(b)
Figure 2. Classification maps of 2 example images from two different patients, achieved using the dSGLD features. The classification done by the radiologist on the superimposed transparency is shown here in yellow.

A classification map was generated for each mammogram and the classification accuracy was computed and displayed using a confusion matrix [17] as shown in table 1. The elements of the confusion matrix, \( N_{i,j} \), give the number of times that a class, \( i \), object was assigned to class, \( j \). The diagonal elements indicate correct classifications, and the off-diagonal elements represent the errors made by the classifier. If the matrix is not normalized, the sum of rows, \( i \), is the total number of elements of class, \( i \), that actually appeared in the data set. The percentage of classification accuracy is given by dividing the sum of the un-normalized diagonal elements by the total sum of the \( i \) rows.

**DISCUSSION**

The minimum distance classifier was used in this study instead of the K-means algorithm to classify the unseen mammograms, because the end cluster centres using the K-means did not fall within the breast tissue types under investigation. Therefore the nearest neighbour algorithm was successfully used to classify the whole mammograms. The use of a two-step classification strategy resolved the confusion between the pectoral muscle and dense tissues, and hence improved the overall classification accuracy. The classification results showed that dSGLD features performed significantly better than the SGLD features in respect of all tissue types as shown in Table 1. Overall, the dSGLD and SGLD features showed an average classification accuracy of 80% and 71%, respectively.

The tissue classification accuracy using dSGLD features for the different tissue types can be summarized as follows: The classification accuracy for fat tissues was the most successful 93%, because the image texture within fat is very different from the rest of the tissues, in that it has very few values off the diagonal (i.e. \( l > 0 \)) of the SGLD matrix as shown in Figure 2. This indicates that fatty tissue has less small scale structure. In addition, the radiologists tended to be more confident in classifying this type of tissue.

The classification accuracy of glandular tissues 77% showed that some 10% of the glandular tissues were misclassified as dense, or connective tissues and that mostly occurred within transition regions i.e. close to other tissue types. Generally this occurred because it is difficult for
the radiologist to visually find a clear border between these types of tissues due to the gradual variation in intensity in the mammogram. Therefore this tended to cause uncertainty in the ground truth classification which was then interpreted as a misclassification. The classification accuracy of dense tissue 84% was improved as the result of using 2-step classification, which resolved the confusion between this class and pectoral muscle, although some 10% of these tissue were misclassified as glandular tissue. Some 4%, tissues in the dense region were considered as unknown tissues “outliers” and may be calcifications.

The classification accuracy for connective tissue was the worst 57%. The connective tissue in the mammograms appeared to consist of thin linear structures. Therefore when the connective tissues were found in a window dominated by fat or glandular tissue they tended to be misclassified as one of those tissues. This was because the connective tissues in such windows were most probably represented by fewer entries in the SGLD matrix than the other tissue types.

Generally, the radiologist classified the selected regions according to the predominant tissue types. Usually the radiologist characterized the breast tissues as mixed tissue types for example, fibro-glandular or fibro-fatty tissues. For the purpose of this study the radiologists were asked to select regions having a single tissue type, but due to the mixed tissue nature of the mammogram images it was, in principle, impossible to always select such unique regions since there was always superposition of tissue types. Therefore, within the region that was demarcated by the radiologist as a single tissue type, there were some sub-images that belonged to other tissue types. These sub-images were usually classified differently by the classifier, and in respect to the ground truth were calculated as being misclassified.

CONCLUSIONS
A comparison of the classification accuracies using the full SGLD matrix and a subset (dSGLD) of these elements showed that the dSGLD technique was superior, and was computationally less intense (30s vs. 1800s) since fewer matrix elements were needed for the calculation. The dSGLD features look promising as classifiers for normal breast tissue composition and they could possibly be used to. Thus we conclude that the classification of the different breast tissues using these new dSGLD features extracted from a subset of the SGLD matrix has advantages over using the full SGLD matrix. Furthermore, this dSGLD technique could be used to classify breast tissues as normal or malignant by further analysis of the dense and glandular tissues. Malignant tumours often appear dense thus making differentiation from dense normal tissues difficult. This method may allow differentiation on grounds of the texture of the tissues rather than simply relying on their greyscale intensity in the radiograph.
REFERENCES