INT J COMPUT COMMUN, ISSN 1841-9836
8(4):499-513, August, 2013.

Combining Feature Methods for Content-Based Classification of Mammogram Images

K. Chikamai, S. Viriri, J-R. Tapamo

Keith Chikamai
Serestina Viriri*
School of Mathematics, Statistics and Computer Science
University of KwaZulu-Natal, Westville Campus
Durban, South Africa
viriris@ukzn.ac.za
*Corresponding author

Jules R. Tapamo
School of Engineering
University of KwaZulu-Natal, Howard College
Durban, South Africa
tapamoj@ukzn.ac.za

Abstract: Breast cancer is among the leading cause of death among females. Studies show that early detection allows for a better prognosis. Mammography is one of the successful ways for early detection of breast cancer. It mostly involves manual reading of mammograms, a process that is difficult and error-prone. This paper discusses a classification model for mammograms based on microcalcification characteristics, as a way of helping radiologists make quick and accurate diagnostic decisions by availing to them similar past cases. The images are pre-processed by Gaussian smoothing and median filtering with 5×5 and 3×3 kernels respectively. Gabor and Haralick features are then extracted to form the image signatures over which similarity measurements are made. Experimental results show an average precision value between 0.5 and 0.61 using Haralick features, 0.49 and 0.57 using Gabor features, and 0.51 and 0.78 using combination of Gabor and Haralick features.
Keywords: Mammogram, Classification, Gabor filters, Grey Level Co-occurrence Matrix, Haralick Features.

1 Introduction

Breast cancer is a high-mortality disease, and one of the leading causes of death among women [1] . If detected early, this disease is manageable, and can be cured in some instances. Mammography is one of the methods used to detect early signs of cancer, and has proven to be very effective [23]. It involves generating images of the breast through X-ray photography, enabling the visualization of the internal breast structure for analysis that can expose any abnormality. Traditionally, mammogram analysis has been manually done by radiologists through visual inspection. The amount of medical images being generated is increasing exponentially. According to Geneva radiology, at Geneva University and Hospitals, images in excess of 30000 are being produced daily [1]. A large image database increases the referential space, providing a solid foundation for solving new cases easily. However, the large size increases the time needed for processing of the images. Speed of decision making is important, especially in medical diagnosis.

Computer Aided Diagnosis (CADx) employs techniques in image processing and artificial intelligence to assist pathologists arrive at objective conclusions about a given image [3]. It is commonly used for identification of suspicious regions in a mammogram, as well as for determination of malignancy. A major part of malignancy determination involves extracting and computing features that are used to characterise the image. Accurate characterization of these

features is important to the overall performance the CAD system and can significantly reduce the rate of unnecessary biopsies. Researchers have explored feature extraction methods to characterize breast pathology that include: morphological [3], wavelets [12], fractal and histogram-based measures. However, efficient and accurate retrieval of images based on their content as a field of computer vision is still an open problem [2, 4]. This research work implements the Gabor filter and the Grey Level Co-occurrence Matrix (GLCM) in an attempt to enable efficient and accurate retrieval of mammogram images containing microcalcifications, as described by a pathologist.

The rest of the paper is organized as follows. Section 2 looks at related relevant work in the literature. Section 3 discusses the proposed system including the features used. Experimental setup and results are discussed in section 4 and the paper is concluded in section 5.

2 Related work

Mammogram images lack color information, and usually exhibit low intensity ranges as well as noise occlusions. Overlying vessels and tissues also present a lot of challenges to the detection of malignant objects. This restricts the type of applicable features to those that exploit shape and textural characteristics of objects, with the requirement that these features be stable and robust against the mentioned limitations.

Researchers have attempted Haralick features for the determination of malignancy in mammograms. Hamid et al. [12] attempted a comparison between wavelet, Haralick and shape features for classification of benign and malignant tumors in mammograms. Pre-processing phase included segmentation using adaptive filtering banks described in [14]. Martins et al. [13] combined Haralick features with shape features as input to the K-means and SVM classifier, achieving considerable success rate of 85%.

Muller et al. [4] look at developments in content-based image retrieval (CBIR) in medical domain and present some future promising research directions. The authors note that speed as an evaluation parameter is rarely mentioned yet is important for an interactive system. They also propose that performance comparison for different feature sets needs to be done to identify well performing visual features and their optimal applications. In pointing out future research directions, the study revealed that availability of good quality features could increase accuracy in data mining and related applications. Specialization is also proposed as a means of including the domain knowledge as a measure of improving accuracy.

Wei et al. [15] analytically look at the potential of CBIR in Medical image database retrieval, and discuss the benefits and feasibility of applying it, or extending the current techniques in order to apply them to daily medical practice. They review the limitations of the current non-CBIR approach, as well as obstacles of the application of CBIR to medical image retrieval. They propose a textural analysis approach based on Grey-level Co-Occurrence Matrices for CBIR in Mammography as a case study. The method involves two stages: image analysis and image retrieval. Image analysis determines the discriminating textural features that best act as descriptors for the image, later to be used for the subsequent image retrieval process. Twelve GLCMs were constructed in four directions at three distances. Eleven Haralick features [11] were then calculated for the 12 GLCMs giving a total of 132 features for each Region of Interest (ROI). The $L2$ norm was used for similarity measurement, with the smallest distance indicating most similarity. A total of 329 ROIs were used for evaluation from images sourced from the Mammographic Images Analysis Society (MIAS). Precision and recall were used to test accuracy, with the system achieving 51% and 19% as the highest scores respectively. Both these values were scored at a GLCM distance of 5. The study identified a number of research issues, which include: semantic gap, systems integration, usability and performance evaluation. The major problems identified with current retrieval systems include subjectivity, financial and time costliness and

inefficiency of image data representation. Obstacles to CBIR application include image noise and many image formats available.

3 General Structure of Proposed Model

The proposed model follows the typical CBIR model (Figure 1). It is generally composed of two stages: off-line feature extraction and on-line image retrieval. Off-line feature extraction involves extraction of features from each of the database images and their storage into the feature domain space. Features are stored in feature vectors which are descriptors of their corresponding images. During on-line image retrieval, a user supplies a query example image whose features are also extracted and used by subsequent algorithms against the database features.

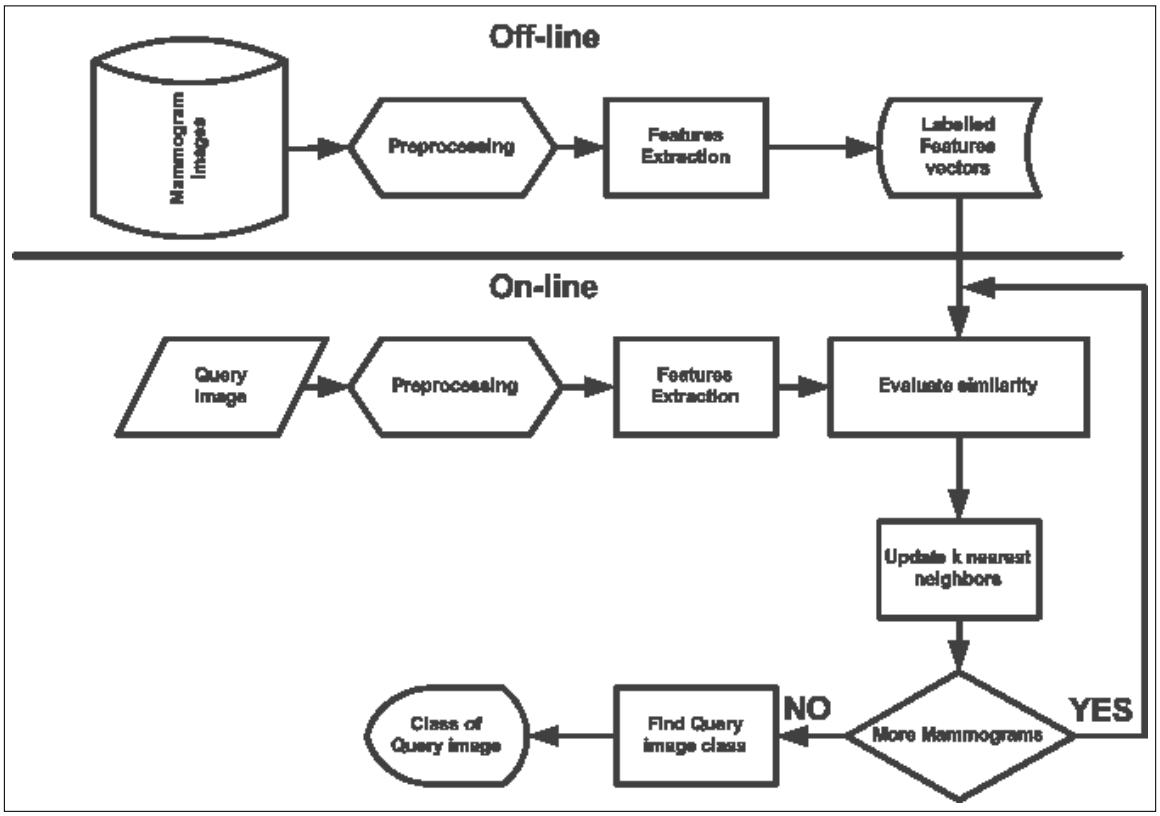

Figure 1: Different steps of a content based classification of a candidate mammogram

3.1 Pre-processing

Breast ROI extraction and denoising

The three techniques applied prior to features extraction are: median smoothing, gaussian smoothing, and region growing. Given an image I with R rows and C columns, the preparation of this image for further processing is carried out in the three following phases:

Phase 1:
The image is first processed using a median filter to remove sporadic sharp frequencies that are characteristics of digital noise. This filter has some edge-preserving characteristics since it does not adversely blur edges. Assuming $\mathcal{N}$ is the kernel, made of neighborhood

pixels around the target pixel ,(x_0, y_0), defined as

$$\mathcal{N} = \{(x_{-n}, y_{-m}), (x_{-n+1}, y_{-m}), ..., (x_n, y_m)\} \tag{1}$$

where $n \times m$ is the kernel size. The median filtering of the (x_0, y_0) is calculated as follows,

- Sort $\mathcal{N}$ in the sequence $S = (S_i)_{i=0,1,...,(n-1)\times(m-1)}$,
- Assign the median value of the sorted sequence to the target coordinates i.e. $I(x_0, y_0) = S_{\frac{(n\times m)}{2}}$

In this case, $n = m = 3$

Phase 2: Gaussian smoothing is then applied on the output image using a kernel $g(x, y)$ of dimension 5×5. The kernel used is of the form,

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{x^2+y^2}{2\sigma^2}} \tag{2}$$

Phase 3: After denoising, region growing is applied to detect the breast region of interest and isolate artificial labels. In order to increase the efficacy of the process by insensitizing the algorithm against small random signal perturbations characteristic of noise, the region growing considers blocks of 8 pixels at a time instead of individual pixels.

Local Gradient and Contrast Enhancement

The image is further enhanced using gradient and contrast enhancement techniques [20]. Gradient enhancement increases the intensity of pixels in an adaptive manner. Taking $I(x, y)$ as the intensity function of a 2D image, the gradient at a pixel (x, y) in a mammogram image is given by,

$$g(x, y) = \frac{1}{n \times n} \sum_{i=-\frac{m}{2}}^{\frac{m}{2}} \sum_{j=-\frac{n}{2}}^{\frac{n}{2}} |I(x + i, y + j) - I(x, y)| \forall x, y \in S \tag{3}$$

The calculated gradient values are then added to the original image to give the gradient enhanced image $I\prime(x, y)$,

$$I\prime(x, y) = I(x, y) + g(x, y) \tag{4}$$

where

$S = \{(x, y) \mid 0 \leq x \leq C - 1,\ 0 \leq y \leq R - 1\}$ is the set of all image pixels

m, n - are the vertical and horizontal spatial dimensions of the kernel m and n determine the extent over which the gradient value is calculated and by implication, the size of objects that are enhanced. A square kernel is used in this work i.e. $m = n = 3$. This technique increases the intensity of pixels relative to the gradient of their local neighborhood. Those areas presenting a higher gradient will thus have their intensities increased more, as determined by the kernel size. The kernel size is intuitively chosen to approximate the spatial extent of microcalcifications in order to enhance their gradient.

Contrast enhancement uses the mean of a region to alter its pixels intensities. The mean of a pixel neighborhood is iteratively calculated as follows,

$$\mu^k(x, y) = \frac{1}{m \times n} \sum_{i=-\frac{m}{2}}^{\frac{m}{2}} \sum_{j=-\frac{n}{2}}^{\frac{n}{2}} \mu^{k-1} I(x + i, y + j) \forall x, y \in \sum * \tag{5}$$

The contrast enhanced image is represented as follows,

$$I''(x,y) = \frac{\mu^k(x,y)}{L-1} I'(x,y) \tag{6}$$

where k specifies the iterations over which the mean is calculated, μ^k is the mean at the kth iteration, L is the highest intensity in the image * the other variables are defined as for gradient enhancement.

While gradient enhancement increases the intensities of high gradient areas without affecting the rest, contrast enhancement diminishes the effect of low contrast areas by lowering their intensity values. The neighborhood determines the objects (by size) that are enhanced. The number of iterations, k, determines how much the original signal is attenuated. A value of $k = 2$ is empirically chosen for the experimental runs in this work. Similar to the gradient kernel, a square mask of dimension 9 (i.e. $m = n = 3$) is used for the contrast kernel. These filters are useful in reducing the effect of the monotonous pelvic muscle which mimics the gradient levels of microcalcifications.

3.2 Feature extraction

The Gabor filter and Haralick features are used for textural analysis of the pre-processed image. The Gabor filtering is an intermediate stage, with the first and second moments being calculated from the Gabor filtered image to give the final features. The techniques and the specific parameters used are discussed in the following sections.

Gabor Filtering

Gabor filter is a transform function related to the Fourier transform which can be used to convey spatial information in addition to frequency properties of a signal. It is commonly applied as a band-pass filter in signal processing where it is used to determine the sinusoidal frequency and phase content of local sections of a time varying input signal, and has been found useful in image compression. Among other useful properties, the Gabor filter has been found to better minimize the conjoint time-frequency information resolution of a signal [4].

A Gabor filtering is obtained by multiplying a complex sinusoidal plane wave of a certain frequency with a Gaussian envelope as follows [5]:

$$g(t) = ke^{j\theta} w(at) s(t) \tag{7}$$

where

$w(t) = e^{-\pi t^2}$ is the Gaussian envelope,

$s(t) = e^{j(2\pi ft)}$ is the sinusoidal function, where f is the frequency of the sinusoidal plane wave, k is the constant, and $e^{j\theta}$ determines the orientation

The strength of the Gabor filter response depends on the filter's congruence with the local signal; where the filter's sensitivity is determined by tuning of the parameters that include: orientation, phase and frequency [5]. Given that Gabor filters are not inherently orthogonal [9], this section determines an optimal set of parameters for designing Gabor filter jets that will detect the desired range of object characteristics with minimum redundancies. Based on the work in [6], this project implements the following 2-D Gabor filter:

$$g_{\lambda\theta\varphi}(x,y) = \frac{1}{\sqrt{2\pi\sigma_x\sigma_y}} e^{-\left(\frac{x'^2}{\sigma_x^2} + \frac{\gamma y'^2}{\sigma_y^2}\right)} cos(2\pi x'w + \varphi) \tag{8}$$

$$x\prime = xcos(\theta) + ysin(\theta) \tag{9}$$

$$y\prime = ycos(\theta) - xsin(\theta) \tag{10}$$

where

λ, θ and φ are configurable input parameters specifying wavelength, orientation and phase, respectively; γ specifies the aspect ratio of the gaussian envelope. Its value is empirically set to 0.5; σ_x and σ_y define the convergence of the gaussian envelope along the $x-$ and $y-$ axes respectively. They are defined as follows,

$$a = \left(\frac{U_h}{U_l}\right)^{\frac{1}{S-1}}$$

$$\sigma_u = \frac{U_h\,(a-1)}{(a+1)\,\sqrt{2log2}}$$

$$\sigma_v = tan\left(\frac{\pi}{2k}\right)\left[U_k - 2log\left(\frac{2\sigma_u^2}{U_h}\right)\right]\left[2log2 - \frac{(2log2)^2\,\sigma_u^2}{U_h^2}\right]^{-\frac{1}{2}}$$

$$\sigma_x = \frac{1}{2\pi\sigma_x}, \qquad \sigma_y = \frac{1}{2\pi\sigma_y}$$

The spatial width of the filter (FW) is then linearly scaled from the derived standard deviation as follows,

$$FW = 4\sigma + 0.5 \tag{11}$$

This filter width calculation is empirically established to give a good compromise since it does not greatly affect border pixels. The pre-processed input image $I(x, y)$ is convolved with Gabor filter $g(s, t)$ to give the response image $r(x, y)$ as shown in Equation (12),

$$r(x,y) = \int\int_{(x,y)} I(s,t)g(x-s, y-t)\delta s\delta t \tag{12}$$

This implementation (Equation 12) is computationally expensive in the spatial domain, making it impractical for large input images and Gabor kernels. This study thus takes advantage of the convolution theorem to implement filtering in the frequency domain as defined in Equation 14. Since convolution in the spatial domain is equivalent to multiplication in the frequency domain, the computational complexity is reduced using the latter method [21]. It uses the comparatively optimal FFTW library to carry out the fourier-spatial transformations.

$$\mathfrak{F}\left(r(x,y)\right) = \mathfrak{F}\left\{I(x,y)\right\}\mathfrak{F}\left\{g(x,y)\right\} \tag{13}$$

$$g(x,y) = \mathfrak{F}^{-1}\left\{\mathfrak{F}\left\{I(x,y)\right\}\mathfrak{F}\left\{g(x,y)\right\}\right\} \tag{14}$$

This work considers a set of Gabor filters configured with four orientations: 0, 45, 90 and 135. These values are calculated according to Equation 15 considering recommendations in [[7, 8]].

$$\theta_k = \frac{k\pi}{n},\ k = \{0, \ldots, n-1\} \tag{15}$$

where n is the number of orientations, k represents the kth orientation.

This method ensures equidistant spacing in the orientation field. Furthermore, the orientation space is chosen from the range $\theta_k \in [0, \pi]$, which provides sufficient coverage since it has been established that response values in the range $[\pi, 2\pi]$ only differ from those in $[0, \pi]$ by phase shift [7, 8].

GLCM and Haralick features

In this work, Haralick features are extracted to describe the Mammogram's textural characteristics. Textural characteristics are described by patterns of pixel intensities [11]. Practically, these intensities are described by a distance-angular relationship model using a Grey Level Co-occurrence Matrix (GLCM) as proposed by Haralick et al. [11]. GLCMs are second-order statistics that define relationships between distinct tonal intensities by measuring the frequency with which they occur together at certain directions (θ) and distances d, and fall under statistical textural classification approaches.

Grey Level Co-occurrence Matrix feature method is also used to model grey level dependencies of mammogram images. Similar to the Gabor wavelets, set of matrices are defined over four directions: 0, 45, 90 and 135, at a distance d and is represented as $P(i, j, \theta, d)$. A subset of nine Haralick features [11] are calculated over the GLCM to describe the mammograms' textural characteristics. For notational convenience, lets denote $P(i, j)$ as the probability of i occurring alongside j, the Haralick features are then defined as follows:

$$Energy = \sqrt{\sum_{i,j} P(i,j)^2} \tag{16}$$

$$Entropy = -\sum_{i,j} P(i,j) log(P(i,j)) \tag{17}$$

$$Contrast = \sum_{i,j} P_{i,j}(i-j)^2 \tag{18}$$

$$Homogeneity = \sum_{i,j} \frac{P_{i,j}}{1+(i-j)^2} \tag{19}$$

$$\text{Max prob} = MAX(P(i,j)) \tag{20}$$

$$Correlation = \sum_{i,j} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{\sigma_i^2 \sigma_j^2}} \right] \tag{21}$$

$$Dissimilarity = \sum_{i,j} P_{i,j}|i-j| \tag{22}$$

$$idm = \sum_{i,j} \frac{1}{1+(i-j)^2} P(i,j) \tag{23}$$

3.3 Classification of a Candidate Mammogram

The CBIR system designed, returns results ranked in order of relevance to the query image. The number of returned images (represented by k in this context) impacts on accuracy or precision of the system if factored during system evaluation. The k Nearest Neighbor (k-NN) classifier is used for classification; we used the version described in [9, 22]. In our context there are two classes M_0, and M_1:

Algorithm 1 :K-NN Algorithm to Classify a mammogram. Given a mammogram c_m, classify c_m based on a database of mammogram grouped into two clusters; one with mammograms containing mammograms microcalcifications and the other one with the ones without microcalcifications.

Require: M, c_m, k ▷ M, is the set of mammograms in the database (each mammogram is labeled belonging to M_0 or M_1) and c_m is the candidate mammogram to be classified. k is the number of nearest neighbors to be considered.

Ensure: $m \in M_l$, ▷ means M_l $(l = 0, 1)$ is the class to which belongs the mammogram c_m

1: Let S be a sequence k elements, ▷ s_i will be the i^{th} element of S
2: $S = \emptyset$ ▷ Initialize S to an empty sequence
3: **for all** element $y \in M$ **do**
4: **insert**(S, y, c_m) ▷ insert y in S, in increasing order of distance between y and c_m,
5: **end for**
6: $l = \arg \max_{v \in \{0,1\}} \sum_{i=1}^{k} I_{M_v}(s_i)$ ▷ This means $c_m \in M_l$

- M_0: class of mammograms without microcalcifications
- M_1: class of mammograms with microcalcifications

$insert(S, y, c_m)$ inserts y in S, in increasing order of distance between y and c_m. Given a set A, $I_A(.)$ is and indicator function defined as follows:
$I_A(x) = 1$ if $x \in A$ and $I_A(x) = 0$ otherwise.

The classifier in **Algorithm 1** finds the minimum distance between the given query vector c_m and all mammograms in M, builds a sequence S of k vectors representing the mammogramms with the minimum distance. The query vector is finally assigned to the class $M_l(l = 0, 1)$ that has the majority of elements of S.

Experiments are conducted with values of k taken from an set of 5 elements, $\{1, 3, 5, 7, 9\}$.

4 Results and Discussion

4.1 Image dataset

Experimental tests were conducted on a set of 60 images sourced from the Mammographic Image Analysis Society (MIAS) database [10]. This database contains 322 mammogram images of mixed pathologies, and is accompanied by ground truth that has been verified and marked by radiologist. The ground truth gives the severity of the pathology (Malignant/benign) as well as the spatial locality and extent of the pathology. The pathology classes are:

- Calcification
- Well-defined/circumscribed masses
- Spiculated masses
- Other, ill-defined masses
- Architectural distortion
- Asymmetry

- Normal

The breast tissue of each mammogram has been classified depending on density. The density of a breast tissue alludes to the amount of fat present in that tissue. Fatty tissues appear as relatively darker areas in mammograms. In order to reduce the amount of fatty tissue, the mammograms are classified under any of these three types: Fatty, Fatty-glandular and Dense glandular.

All dataset images have been quantized to 256 grey levels, and were digitized at a resolution of 200 microns. They are padded and clipped to occupy a standard size of 1024×1024 pixels.

4.2 Experimental setup

Sixteen images were selected from the database to form the query image set. These images were taken from both classes of pathology, i.e., normal and malignant. Normal images in this experiment were defined as those images not containing Microcalcifications. This definition covers images diagnosed as positive for other malignancy indicators such as circumscribed masses and asymmetries. Results were then collected for each round for every query image. For generalized results, precision values calculated are averaged over all query images instead of one. The query process was repeated ten times (ten iterations) using a randomized set of 8 images from the normal class, the average precision values obtained in each round were then averaged over the ten iterations to give the statistical base for reporting. This process was done for every value of $k \in [1, 3, 5, 7, 9]$.

4.3 Performance Metric

The average precision curve [18] was used to evaluate the performance of the system. Precision gives the general classification performance of the system. It measures the ability to correctly classify both sample sets. Sixteen images from both classes are used as query images, and results collected after every round as explained in section 4.2. For every returned result set, precision is calculated as follows,

$Precision = \frac{R}{k}$, where R is the number of accurate predictions and k the number of neighbors

Algorithm 2 is used to compute precisions. The average of precision values for both sets of query images is then taken as the precision value for the round. The precision values are then used to sketch the precision curve for diagrammatic representation. The precision value ranges from a minimum of 0 to a maximum of 1. A high precision value implies that the system has a commensurately high ability of correctly classifying a given sample.

4.4 Discussion

The first results of the experimental runs are given in Figures 3, 4 and 5, show the Haralick results for single, double and combined class query image sets, respectively. The figures 6, 7 and 8, show the Gabor filter results for single, double and combined class query image sets, respectively. The figures 9, 10 and 11 show the results for single, double and combined class query image sets, respectively, using combined Gabor and Haralick features.

The first results (Figures 3, 4 and 5), show precision values obtained by querying the database using 8 images randomly selected from the "Normal" class. The querying process (section 4.2) is evaluated over five rounds. The lowest precision value of 0.71 is scored at distance $k = 3$, with the highest value of 0.88 scored at the distance $k = 1$. The system gives a low score for images identified positive for microcalcifications . The highest score of 0.375 is recorded at distance $k = 1$, and the lowest score of 0.13 scored at distance $k = 9$. For mixed class query images (Figure 4), $1-$Nearest Neighbor gives the best precision score at 0.69, with the $9-$Nearest

Algorithm 2 Retrieval performance benchmarking

$Query\_Set \leftarrow getRandomImages(Database)$
for all kNN *distances* k **do**
  $Precision(k) \leftarrow 0$
  for $i = 1$ **TO** $numOfIterations$ **do**
    for all $Query\_image\ q\ in\ Query\_Set$ **do**
      $Result \leftarrow getNearestNeighbors(q)$
      $Precision(k) = Precision(k) + getPrecision(Result)$
    end for
    $Precision(k) \leftarrow Precision(k)/no.ofqueryimages$
  end for
  $Precision(k) = Precision(k)/no.ofIterations$
end for

Figure 2: Query result using Normal class images

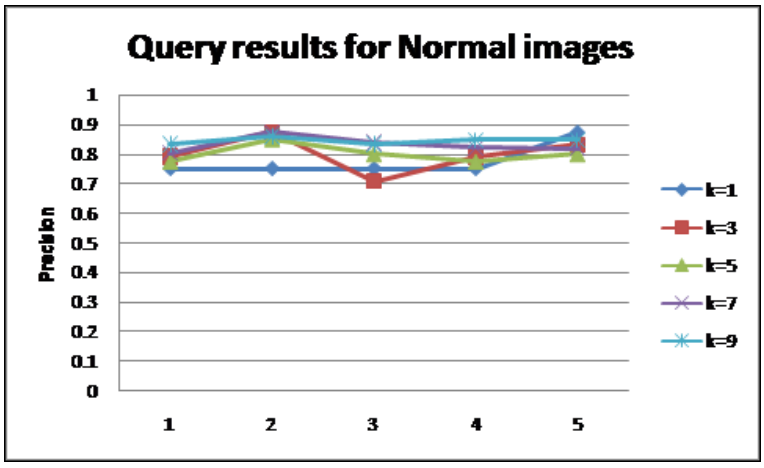

Figure 3: Query result using Normal class images

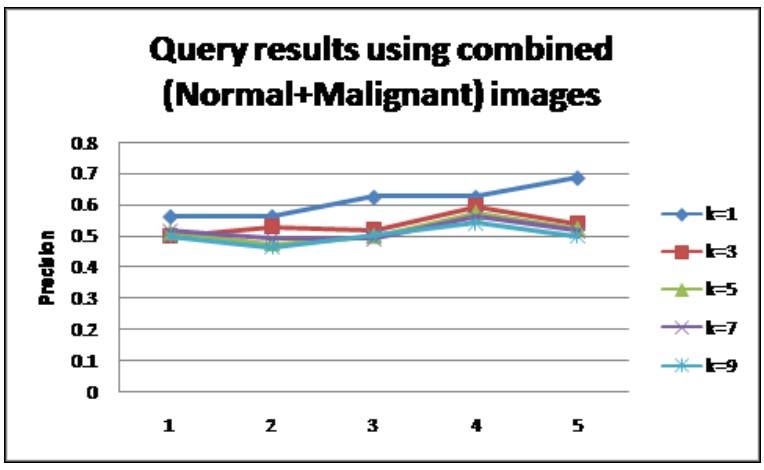

Figure 4: Query results using mixed class images

Neighbor giving the lowest score at 0.47. Overally, the performance degrades with an increasing value of k (Figure 5). The highest value scored is 0.61 at $k = 1$, with the lowest value of 0.50

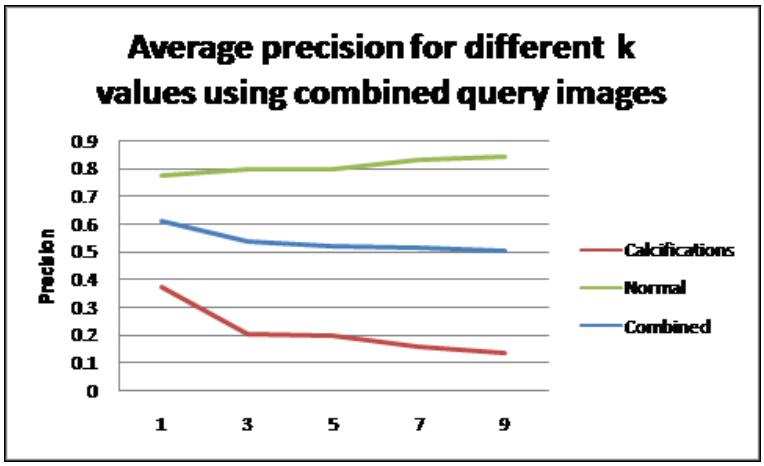

Figure 5: Average Precision for Combined class query image set

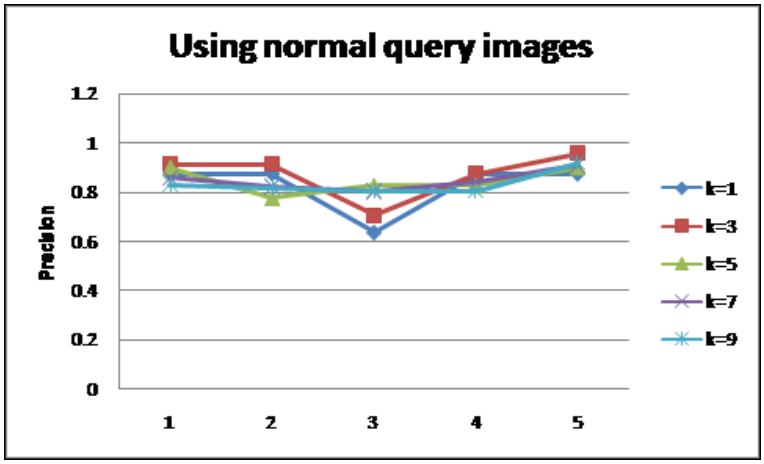

Figure 6: Query result using Normal class images

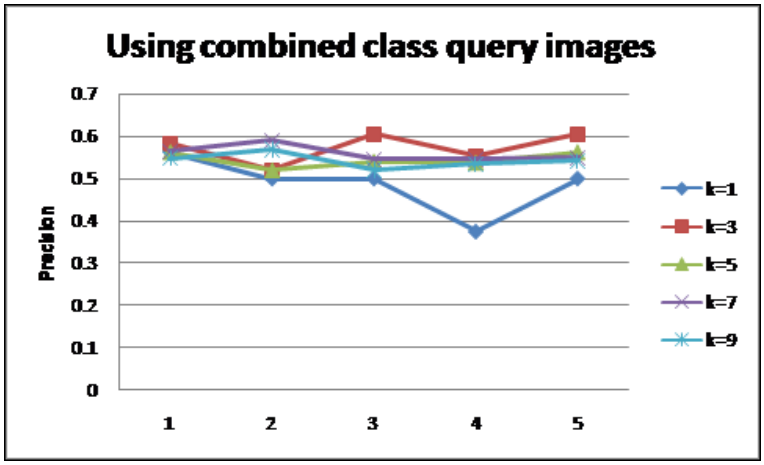

Figure 7: Query results using mixed class images

being scored at $k = 9$.

Compared to Haralick features, the Gabor feature set gives slightly higher average values

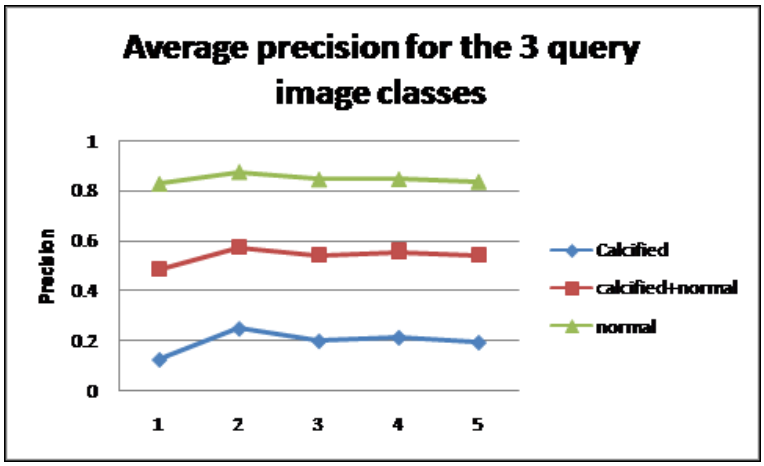

Figure 8: Average Precision for Combined class query image set

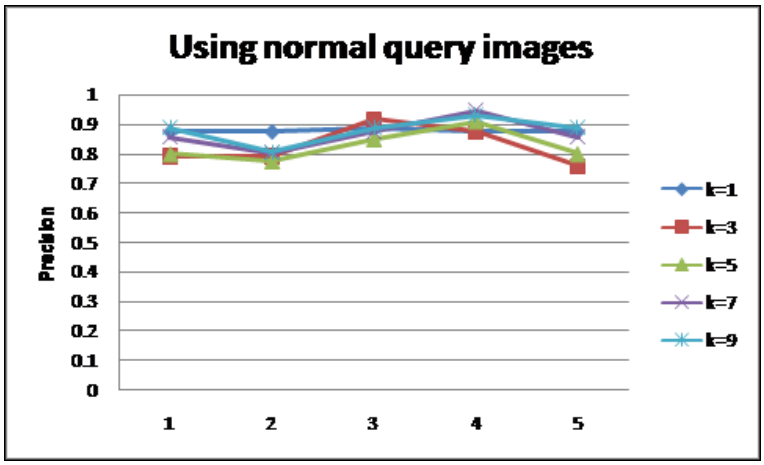

Figure 9: Query result using Normal class images

Figure 10: Query results using mixed class images

(Figures 6, 7 and 8) for queries involving benign classified images with a high score of 0.85 (considering all values of k). However, the Gabor vector gives relatively lower high scores for the

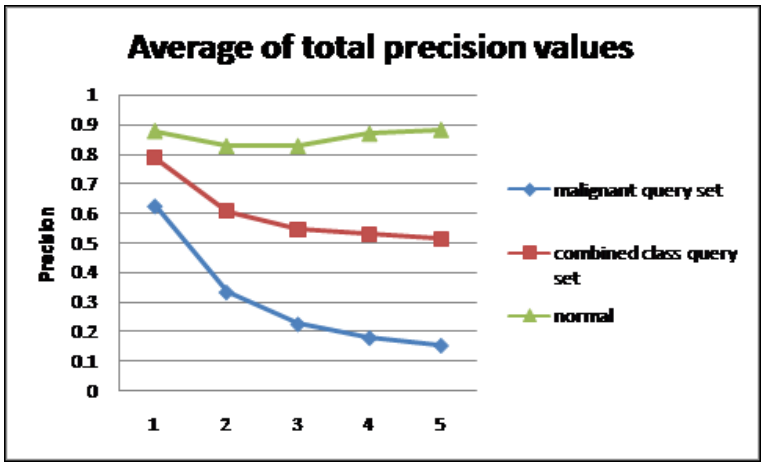

Figure 11: Average Precision for Combined class query image set

calcification and combined class query image set, at 0.25 and 0.57 respectively. It however gives consistently better average results considering all values of k for all three query classes.

Comparison of retrieval performance was done considering a mixed Gabor and Haralick feature set (Figures 9, 10 and 11), which gave mixed results for various values of k. The combined feature set gave the highest score for queries involving normal classified images, with a score of 0.88 at $k = 9$ (Figure 9). It also registered the highest scores for $k \in [1, 7]$, with the Gabor vector giving the highest values for the remaining values of k. This set also gave high values for the calcification and mixed class query images at 0.625 and 0.79 respectively, both values attained at $k = 1$. Summarily, the combined feature set of Gabor and Haralick features enhances retrieval performance for all classes of query images. The best consistent performance is achieved at $k = 1$ for all query classes.

For comparison, Wei et al. [19] implement a GLCM-based mammogram retrieval system for comparison with their algorithm. Regions of interest (ROIs) are cropped from mammogram images of multiple pathologies. The images are then Gabor-filtered before calculation of Haralick features. The GLCM matrix is calculated over three distances and four orientations. Their average precision values range between 0.33 and 0.64. Our system proves to be more discriminating towards images not containing microcalcifications than for those containing microcalcifications. A possible explanation is that a lot of unnecessary breast information is being included for similarity calculations. This means that the algorithm needs to be enhanced more to reduce the impact of non-calcification regions. The high dimension of features might also have a negative effect on the accuracy of the algorithm by introducing redundancies. Our precision value is however not far-off the one obtained in [19]. This adds to the fact that the proposed model automates ROI selection.

5 Conclusions and Future Work

Mammography allows the detection of breast cancer in its early stages, which makes possible early remedial measures that can reduce the high mortality rates associated with the disease. This paper discussed a content-based classification model for mammogram images, with the objective of availing a second opinion to a radiologist for reference during diagnosis. It implemented the Gabor filter and Haralick features for textural analysis and description for similarity assessment. This work evaluated Haralick features at five distances, $k \in [1, 3, 5, 7, 9]$ and four orientations East, North, South and West. The best value is attained using a combined Gabor and Haralick

feature set, with a score of 0.79 at the distance of $k = 1$, with lowest value as 0.49 at the distance $k = 9$ using Gabor features only. The moderate precision value could be attributed to the impact of non-calcification breast areas, as well as redundant and less discriminating features. Work is underway to remove redundant features as well as those features that do not discriminate well with respect to microcalcifications.

Bibliography

[1] Rangayyan, R.M. (2005); Breast Cancer and Mammography, *Biomedical Image Analysis, Springer-verlag*: 22-27.

[2] Chu, K.C. et al (1996); Recent Trends in US Breast Cancer Incidence, Survival, and Mortality rates, *Journal of the National Cancer Institute*, ISSN 1460-2105, 88(21): 1571-1579.

[3] Oliver, A.; Freixenet, J.; Marti, R.; Zwiggelaar, R. (2006); A Comparison of Breast Tissue Classification Techniques, *MICCAI, CRC Press*: 872-879.

[4] Hamid, S.Z.; Farshid, R.R.; Siamak, P.N. (2004); Comparison of Multiwavelet, Wavelet, Haralick, and Shape Features for Microcalcification Classification in Mammograms, *Pattern Recognition* 37: 1973-1986.

[5] Li, C.T.; Wei, C.H; Wei, C.H.; Li, C. (2005); A Content-based Approach to Medical Image Database Retrieval, *Database Modeling for Industrial Data Management: Emerging Technologies and Applications* 10(6): 681-685.

[6] Muller, H.; Michoux, N.; Bandon, D.; Geissbuhler, A. (2004); A Review of Content-based Image Retrieval Systems, Medical Applications-clinical Benefits and Future Directions, *Int J Med Inform*: 1-23.

[7] Miyamoto, E.; Merryman, T.; Fast Calculation of Haralick Texture Features, *Technical Report, Carnegie Mellon University.*

[8] Martins, O.L. et al (2009); Detection of Masses in Digital Mammograms using K-Means and Support Vector Machine, *Electronic Letters on Computer Vision and Image Analysis*, 8(2): 39-50.

[9] Wei, C.H.; Chang-Tsun, L.; Roland, W. (2009); A Content-based Approach to Medical Image Database Retrieval, *Database Technologies: Concepts, Methodologies, Tools, and Applications*: 1062-1083.

[10] Haralick, R.M.; Shanmugan, K.; Dinstein, I. (1973); Textural Features for Image Classification, I*EEE Transactions on Systems, Machine, and Cybernetics SMC*, 3(6): 610-621.

[11] Lee, S.K. et al (2000); A Computer Aided Design Mammography Screening System for Detection and Classification of Microcalcifications, *International Journal of Medical Informatics* 60: 29-57.

[12] Lee, T.S. (1996); Image Representation using 2D Gabor Wavelets, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10: 959-971.

[13] Turner, M.R. (1986); Texture Discrimination by Gabor Functions, *Biological Cybernetics*, 10: 71-82.

[14] Hamamoto, Y. et al (1998); A Gabor Filter-based Method for Recognizing Handwritten Numerals, *Pattern Recognition*, 31(4): 395-400.

[15] Kruizinga, P. et al (1999); Comparison of Texture Features based on Gabor Filters, *ICIAP*: 142-147.

[16] Liu, C.; Wechsler, H. (2002); Gabor Feature based Classification using the Enhanced Fisher Linear Discriminant Model for Face Recognition, *IEEE Trans. Image Processing*, 11(4): 467-476.

[17] Kyrki, V.; Kamarainen, J.; Kalviainen, H. (2004); Simple Gabor Feature Space for Invariant Object Recognition, *Pattern Recognition Letters*, 10: 311-318.

[18] Dimai, A. (1999); Rotation Invariant Texture Description using General Moment Invariants and Gabor Filters, *Proc. of the 11th Scandinavian conference on Image analysis*: 391-398.

[19] Steinbach, M.; Tan, P.N. (2009); KNN: k-Nearest Neighors, Data Mining and Knowledge Discovery, *Taylor Francis Group*, Ch. 8: 151-161.

[20] Suckling, J. et al (1994); The Mammographic Image Analysis Society Digital Mammogram Database, Exerpta Medica, *International Congress Series*, 1069: 375-378.

[21] Luo, J.; Nascimento, M.A. (2004); Content-based Sub-Image Retrieval using Relevance Feedback, *Proc. of MMDB*: 2-9.

[22] Wei, C.H.; Li, Y.; Li, C.T. (2007); Effective Extraction of Gabor Features for Adaptive Mammogram Retrieval, *Proc. of ICM*: 1503-1506.