Chapter 2: Theoretical Background and Literature Review

This chapter presents the theoretical background related to a CBIR system and a comprehensive literature review related to the work presented in the thesis.

With the rapid advancement of digital imaging technologies and the use of large volume image databases in various applications, it becomes imperative to build an automatic and an efficient image retrieval system. Content-based image retrieval (CBIR) is most emerging and vivid research area in computer vision, in which unknown query image assigns to the closest possible similar images available in the database. CBIR system for general-purpose as well as mammogram images databases is a highly challenging problem because of the large size of the database, feature indexing, semantic and sensory gap, pre-processing limitations, the difficulty of understanding images, searching and browsing issues.

Feature is the property of an object, which can discriminate an object from others. Usually humans use colour, shape and texture to analyse and recollect the contents of an image [7, 24-26]. Also, CBIR systems use these features for the indexing of images, which are archived along with the images by performing similarity match between query image features with those in the database. Therefore, it is natural to use features based on these attributes for image retrieval.

2.1 Features Extraction for CBIR system

2.1.1 Colour Features

Colour is one of the most widely used visual features in content-based image retrieval [24]. While we can perceive only a limited number of gray levels, our eyes are able to distinguish thousands of colours and a computer can represent even millions of distinguishable colours in practice. Colour has been successfully applied to retrieve images, because it has very strong correlations with the underlying objects in an image. Moreover, colour feature is robust to background complications, scaling, orientation, perspective, and size of an image. Although we can use any colour space for computation of a colour histogram HSV (hue, saturation, value), HLS (hue, lightness, saturation), and CIE colour spaces (such as CIELAB, CIELUV) have been found to produce better results as compared to the RGB space. Since these colour spaces are visually (or perceptually) uniform compared to the RGB, they are found to be more effective to measure colour similarities between images [27]. There are variants of colour features are used in literature, some introduction of few are as follow:

- *Colour Histogram*

Colour histogram is one of the most important descriptor used in content-based image retrieval; which shows, how many pixels in an image are of a particular colour [28]. Colour histogram is represented as bar chart, where each bar (bin) represents a particular colour of the colour space being used. For the purpose of saving time, here we have reduced the number of bins through quantization, by taking colours that are very similar to each other and putting them in the same bin.

For an M×N image I, the colours in that image are quantized to Q_1, Q_2…Q_{32}. The colour histogram H (I) = [H_1, H_2, …, H_{32}], where H_i represents the number of pixels in colour Q_i. The colour histogram also represents the possibility of any pixel, in image I, that in colour Q_i.

$$Probality(Prob \in Q_i) = \frac{H_i}{M \times N} \tag{2.1}$$

- *Colour Coherence Vector*

One problem with the colour histogram based similarity measure approach is that the global colour distribution doesn't reflect the spatial distribution of the colour pixels locally in the image. This cannot distinguish whether a particular colour is sparsely scattered all over the image or it appears in a single large region in the image. The colour coherence vector-based [29] approach was designed to accommodate the information of spatial colour into the colour histogram. Here we can classify each pixel in an image, based on whether it belongs to a large uniform region.

- *Colour Moments*

Colour moment is a compact representation of colour features to discriminate a colour image. It has been shown that most of the colour distribution information is captured by the three low-order moments [30]. These are mean, standard deviation and skewness, reflect the average colour value in the image, colour deviation from the mean, and degree of asymmetry in the distribution, respectively.

- *Colour Correlogram*

The weak point of the histogram method is lack of space information in colour. Colour correlogram is technique proposed to integrate spatial information with colour histograms. For each pixel in the image, the correlogram approach needs to go through all the neighbours of that pixel. So the colour correlogram shows how the spatial autocorrelation of colour changes with distance [31].

- *Fuzzy Colour Histogram (FCH)*

Colour histogram considers neither the colour similarity across different bins nor the colour dissimilarity in the same bin. Therefore, it is sensitive to noisy interference such as illumination changes and quantization errors. To address these concerns, a new colour histogram representation, called fuzzy colour histogram (FCH), by considering the colour similarity of each pixel's colour associated to all the histogram bins through fuzzy-set membership function [32]. This feature is robust to quantization error and changes in light intensity.

- *Colour Difference Histogram (CDH)*

Most of the existing histogram techniques merely count the number or frequency of pixels. However, the unique characteristic of CDHs is that they count the perceptually uniform colour difference between two points under different backgrounds with regard to colours and edge orientations in $L^*a^*b^*$ colour space [33]. This method pays more attention to colour, edge orientation and perceptually uniform colour differences, and encodes colour, orientation and perceptually uniform colour difference via feature representation in a similar manner to the human visual system.

- *Colour Layout Descriptors*

A colour layout descriptor (CLD) is designed to capture the spatial distribution of colour in an image [34]. The feature extraction process consists of two parts; grid based representative colour selection and discrete cosine transform (DCT) with quantization.

- *Chromaticity Moments*

This feature has resolve the issue of long histogram, only a small number of features, called chromaticity moments, are required to capture the spectral content (chrominance) of an image[35]. Chromaticity moments are characterized by their two dimensional shape and two dimensional distribution.

2.1.2 Texture Feature and Key Point's Detector

Texture is a very interesting image feature that has been used for characterization of images, with application in content-based image retrieval [36]. There is no single formal definition of *texture* in the literature. However, a major characteristic of texture is the repetition of a pattern or patterns over a region in an image. Brief description of some dominating texture features are as follow:

- *Gray level Co-occurrence Matrix (GLCM)*

GLCM is a statistical approach for computing the co-occurrence probability of different combinations of grey levels in an image [37]. The matrix element $G(p, q \mid \Delta x, \Delta y)$ is the relative frequency, where two pixels are separated by a pixel distance (Δx, Δy) within a given neighborhood, one with intensity p and the other with intensity q.

Let $I(x, y)$ is an image with size $M \times N$ *and* gray levels g ranging from 0 to g-1. Then GLCM matrix for an image I, parameterized by an offset $(\Delta x, \Delta y)$ is defined as [29, 35] :

$$G_{\Delta x,\Delta y}(p,q) = \sum_{x=1}^{M}\sum_{y=1}^{N}\begin{cases}1, \text{if } I(x,y) = p \text{ and } I(x+\Delta x, y+\Delta y) = q \\ 0, \quad \text{otherwise}\end{cases} \tag{2.2}$$

- *Tamura Texture*

By observing psychological behaviour of human visual perception, Tamura proposed the texture representation using computational approximations to the three main texture features of: coarseness, contrast, and directionality [38].Where *Coarseness* is the measure of average regions that have the same intensity, *Contrast* is the measure of distinctness of the texture pattern and Directionality is the measure of direction of the grey values within the image.

- *Wavelet Transform based Features*

Wavelet transform is a signal processing technique extensively used in texture analysis & extraction of visual texture features based on multi-resolution decomposition of the images, and representing textures in different scales [39-40]. Wavelet transform, transforms the images into a multi-scale representation with lower computational cost. When we apply discrete wavelet transform (DWT) to the input images, it decomposes the images into four parts (LL, LH, HL and HH). Further, low-low sub-bands are decomposed, and repeat for LL sub band as desired number of decomposition. Statistical features like mean, standard deviation, skewness, kurtosis etc. of the transform coefficients are used as a feature vector.

- *Gabor Filter based Features*

Gabor filter is an example of linear wavelet filters, capturing energy at a specific frequency and a specific direction, and frequently used in many image processing applications such as; synthesis of images, segmentation, edge detection, pattern recognition etc. In all such applications, it is necessary to analyse the spatial frequency parts of an image in a localized manner using a Gaussian envelope [41]. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have justified being appropriate for extracting useful texture features from an image.

A two dimensional Gabor function $g(x, y)$ is defined as:

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + jw(xcos\theta + ysin\theta\right] \qquad (2.3)$$

where space constants σ_x and σ_y define the Gaussian envelope along the X *and* Y-axes, w is modulation frequency *and* θ is orientation.

After applying Gabor filters on the images with different orientation at different scale, we obtain an array of transformed coefficients. Mean square energy, Mean amplitude of these Gabor coefficients are used to represent the homogenous texture feature of the region. But this features is computational intensive as compare to other transform.

- *Local Binary Patterns*

Local binary patterns (LBP) are a computationally light weighted texture features, derived from the local neighbourhood of each pixel in the image [42]. This operator treats each pixel as a center pixel, and calculates the difference of center pixel with neighbourhood pixel, and multiplies it with a binary image window.

Despite these textures, four pioneering texture and edge features have paved the way to the significant advance in content-based visual retrieval on large-scale multimedia database. The first one is the introduction of invariant local visual feature SIFT [43]. SIFT is demonstrated with excellent descriptive and discriminative power to capture visual content in a variety of literature. It can well capture the invariance to rotation and scaling transformation and is robust to illumination change. The second work is the introduction of Edge in the image is considered an important feature to represent the content of the image. In MPEG-7, there is a descriptor for edge distribution in the image. This edge histogram descriptor proposed for MPEG-7 [44] consists only of local edge distribution in the image. That is, since it is important to keep the size of the histogram as small as possible for the efficient storage of the metadata. SURF (Speed Up Robust Feature) is one of the most and popular interest point detector and descriptor which has been published by Bay et al. [45]. It is widely used in most of the computer vision applications. The SURF has been proven to achieve high repeatability and distinctiveness. Other texture and key points descriptors which are actively used in CBIR are Histogram of gradient [46], Binary Robust invariant scalable key points (BRISK)[47], Features from accelerated segment test (FAST) [48] etc.

2.1.3 Shape Features

Shape is another image feature applied in CBIR. Shape can roughly be defined as the description of an object minus its position, orientation and size [3]. Therefore, shape features should be invariant to translation, rotation, and scale, for an effective CBIR, when the arrangement of the objects in the image is not known in advance. To use shape as an image feature, it is essential to segment the image to detect object or region boundaries; and this is a challenge. Techniques for shape characterization can be divided into two categories. The first category is boundary-based, using the outer contour of the shape of an object. The second category is region-based, using the whole shape region of the object. The most prominent representatives of these two categories are Fourier descriptors, chain code, polygon approximation, moment invariants, curvature scale space descriptor, angular radial transform, image moments, and geometric features etc. [49-52].

2.2 Similarity Measures

Generally image descriptors are proposed along with its distance measure to match the descriptors. The most commonly distance measure used and reported in the literature which computes the dissimilarity between descriptors is the Euclidean distance. If the Euclidean distance between two descriptors is less, it means two descriptors are more likely to be similar; hence the corresponding images are also similar. Other common distances are *L1*, *Cosine, Canberra, D1, and Chi-square* etc. [53-56]. The fundamental work of similarity measure is to find the dissimilarity between the descriptors of two images.

Let the descriptors of two images are as follow:

$$F^a = F^a(1)\ F^a(2), \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, F^a(n)$$

$$F^b = F^b(1)\ F^b(2, \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots, F^b(n)$$

where *n* is the dimension of feature vectors. The different distances are defined as follows:

- Euclidean Distance (ED):

$$ED(F^a, F^b) = \left(\sum_{i=1}^{n} (F^a(i) - F^b(i))^2 \right)^{\frac{1}{2}} \tag{2.4}$$

- *L1* Distance:

$$L1(F^a, F^b) = \left(\sum_{i=1}^{n} \left| F^a(i) - F^b(i) \right| \right) \tag{2.5}$$

- The *D1* Distance

$$D1(F^a, F^b) = \left(\sum_{i=1}^{n} \frac{\left|F^a(i) - F^b(i)\right|}{\left|1 + F^a(i) + F^b(i)\right|} \right) \tag{2.6}$$

- Canberra Distance (*CD*):

$$Canberra(F^a, F^b) = \left(\sum_{i=1}^{n} \frac{\left|F^a(i) - F^b(i)\right|}{\left|F^a(i) + F^b(i)\right|} \right) \tag{2.7}$$

- Chi-square Distance

$$Chisq(F^a, F^b) = \left(\frac{1}{2} \sum_{i=1}^{n} \frac{(F^a(i) - F^b(i))}{(F^a(i) + F^b(i))} \right) \tag{2.8}$$

2.3 Literature Review of CBIR System for General Images

A number of general-purpose image search engines have been developed using different variants of colour, texture and shape features with different variants of similarity measures [56-59]. Recent researches on CBIR are also going on the fusion of dominating features and the role of similarity measures. A brief Literature review of general CBIR system is given in Table 2.1

Table 2.1: Brief description of various CBIR systems

Authors	Description	Limitations/ Issues	Database Used	Performance Parameters
Walia and Pal [9]	Used colour difference histogram (CDH) and Angular Radial Transform based features that capture the colour, texture and shape	Extra pre-processing efforts are required, uses exhaustive linear	Wang, VisTex, and OT-Scene Database	Precision(P), Recall (R), and P-R Graph

	information of an image, and finally to make system more effective, they used fusion framework to combines the ranking results	search in the entire database		
Ela *et al.* [23]	Used Multiple SVM classifiers with wavelet features, and reduced the searching time of retrieval	Lack aspects of generalization	Wang	
Liu and Yang [33]	Based on uniform colour difference between colours and edge orientation, which cover a wealthy visual content for retrieval	Searching time and feature dimension issue (256 features)	Wang	Precision, Recall, Accuracy
Malik & Baharudin [54]	Used quantized histogram statistical texture features from the discrete cosine transform (DCT) blocks of the image. For the effective matching various distance metrics are used	Uses traditional searching technique, loss of information up to some extent	Corel-1000	Precision , Recall, and F-measure
Yue *et al.* [60]	Used colour histogram (CH) and texture features based on a co-occurrence matrix to	Feature dimension is high due to use	Wang	Precision, Recall, Accuracy

	form feature vectors. These fused features were analysed by constructing weights of feature vectors, and conveyed that, fused features retrieval brings a better visual feeling than the single feature retrieval	of CH, colour co-occurrence matrix required more time for indexing, further uses exhaustive linear search in the entire database		
Jalab [61]	Used colour layout descriptor (CLD) and Gabor filter, and concluded that, combining the colour and texture features in CBIR systems leads to more enhancing results for image retrieval.	Size of feature vector,performance is not robust & effective, Semantic gap and searching time	Wang	Precision, Recall, P-R graph
Shen *et al.* [62]	Firstly, they segmented an image into regions and further extracted local colour, texture and CENsus Transform hISTogram features for indexing and retrieval	Extra pre-processing time for segmentation, required more searching time	Wang	Precision, Recall, Accuracy

Irtaza *et al.* [63]	Presented Neural Network based architecture for content based image retrieval, where the feature extraction method was based on depth-texture analysis using wavelet packets and Eigen values of Gabor filters. For correct image retrieval, a partial supervised learning scheme was introduced, based on K-nearest neighbours of a query image	Risk of mis-association, semantic association using backpropagation neural networks, and Class finalization through K-nearest neighbours	Wang, Coil	Precision, Recall, Average precision & Recall class wise
Rahimi and Moghaddam [64]	Used Intra class features based on the concept of co-occurrence matrix, Distribution of Colour and inter class features based on dual-tree complex wavelet transform, (SVD), and conceptual segmentation	Lack of generalization in retrieval performance, required more searching time	Wang, VisTex images	Classification rate, Retrieval rate, ROC curve
Liu *et al.* [65]	Retrieval based on new image descriptor (Micro	Complex, feature	Wang	Precision, Recall , and

	Structure Descriptor-MSD), uses edge orientation similarity, and built upon underlying colours in micro-structures	dimension is moderately high (72 features), issue of searching time		P-R curve
Zhang *et al.* [66]	Used physiological structure of human eyes and visual perception based hybrid information descriptors (HIDs), consisting of mutual information descriptors (MIDs) and self-information descriptors (SIDs)	Feature dimension is high (110 features), Issue of searching time, more indexing time required for extraction of MID and SID	Wang (Corel-1000 and 10,000), Oliva and Caltech	Precision, Recall, Average P-R curve
Zhao *et al.* [67]	Based on colour texture and shape features: Used colour distribution entropy (CDE), colour level co-occurrence (CLCM) and invariant moments. The similarity measure matrix is based upon Euclidean distance	More time required for feature indexing due to use of Invariant moments and Colour level co-occurrence (i.e. more than 1.2	Wang database	Precision

		seconds are required for an image of size 256×256), uses linear search		
Mistry *et al.*[68]	Used hybrid feature with various distance measure. i.e. Spatial domain features including colour auto-correlogram, colour moments, HSV histogram features, and frequency domain features like moments using SWT, features using Gabor wavelet transform	Huge dimension of feature vector, Uses Exhaustive linear search in the entire database.	Wang database	Precision
Singh & Kaur [69]	They have introduced fast and efficient image retrieval system based on colour and texture features. The colour features are represented by colour histograms and texture features are represented by block	Required relatively extra feature indexing time and exhaustive linear search in the entire database	Corel-5K, Corel-10K, UKbench and Holidays	Precision

	difference of inverse probabilities and block variation of local correlation coefficients			
Wang *et al.* [70]	Introduced an effective colour image retrieval method based on texture, which uses the colour co-occurrence matrix to extract the texture feature and measure the similarity of two colour images.	Work only for colour dominated images, uses direct similarity between two images tends to search entire database.	Wang	
Wang *et al.* [71]	New and effective colour image retrieval scheme by combining colour, texture and shape information. For colour , texture and shape features, fast colour quantization with clusters merging, steerable filter decomposition, and pseudo-Zernike moments are respectively used	Complex and required extra processing time for feature extraction, uses linear search for retrieval	Wang	Precision

Banerjee *et al.* [72]	Proposed significant clusters points around significant curvature regions using fuzzy set theoretic approach, and extracted invariant colour features. Minimum redundancy-maximum relevance (mRMR) based informative feature selection technique are used to find most relevant features	Feature dimension is high, Extra step (more time) of feature selection has added.	Wang	Precision, Recall, Accuracy
Ealami [73]	Firstly extracted colour and texture features then apply dimensionality reduction, artificial neural network (ANN) and matching strategy. The ANN serves as a classifier, so that the selected features of query image are the input and its output is one of the multi classes that have the largest similarity to the query image	Extra time needed for selecting the features due to training of ANN	CIFAR-10 Wang, Oxford Flowers dataset	Average Precision and Recall between different aspects
Das *et al.*	Used RF for feature	Extra time	Wang	Precision,

[74]	reweighting where weights in the similarity measure are modified using feedback samples as returned by the user	needed for learning of model for Relevance feedback, performance factor		Recall, Average P-R curve
Dubey *et al.* [75]	Proposed brightness invariant colour transformation based on a multi-channel based illumination compensation, generic model for various features, which retrieval performances are preserved in the case of illumination change	Application specific retrieval framework	Phos Natural Illumination database, Two Corel Synthesized Illumination Database	Average precision rate, average recall rate, Retrieval & Overall average P-R, Average P-R Curve
Ganar *et al.* [76]	Based on colour, texture and shape features using direct similarity measure. Global colour histogram and local colour histogram, co-	Required additional feature dimension and time for	Wang database	Precision and Recall

	occurrence matrix and gradient method are used for colour, texture and shape features respectively	indexing, uses linear search		
Afifi and Ashour [77]	Ranklet Transform and the colour feature are used as a visual feature to represent the images. To speed up the retrieval time, images are clustered according to their features using k-means clustering algorithm	Extra Pre-processing step, No any analysis of invariant properties are shown for Ranklet Transform	Wang	Precision and Recall
Alsmadi [78]	Used feature set is composed of colour signature with the shape and colour texture features. Consequently, a novel similarity evaluation using a meta-heuristic algorithm called a memetic algorithm (genetic algorithm with great deluge) is achieved between the features of the query image and the features of the	Feature dimension is high, meta-heuristic approaches take extra pre-processing time for the optimization of results, uses linear search for retrieval	Wang	Precision, Recall , P-R Curve

	database images			
Varish and Pal[79]	In this work, colour feature is extracted from the quantized histograms of Hue (H) and Saturation (S) components while texture feature is extracted from computed gray level co-occurrence matrices (GLCMs) of each sub image of discrete wavelet transform (DWT) of Value (V) component of HSV colour image	Uses exhaustive linear search for retrieval, Additional pre-processing steps which losses the colour information.	Wang	F-Score, Precision, Recall
Chun *et al.* [80]	Based on an efficient combination of multiresolution colour (correlogram) and texture features. The colour and texture features are extracted in multiresolution wavelet domain. The dimension of the combined feature vector is determined at a point	Uses exhaustive linear search for retrieval	Wang	Precision, Recall

	where the retrieval accuracy becomes saturated			
Huang *et al.* [81]	Used colour and texture features. As its colour features, colour moments of the Hue, Saturation and Value (HSV) are used. As its texture features, Gabor texture descriptors are adopted. Users assign the weights to each feature respectively and calculate the similarity with combined features of colour and texture according to normalized Euclidean distance	Uses exhaustive linear search for retrieval	Wang	Precision, Recall
Liu *et al.* [82]	Used LBP based texture and colour histogram based colour features for image retrieval and classification	Limited performance and feature dimension is moderately high, further also uses exhaustive linear search in the	Wang	Precision and Recall, P-R Curve

		entire database		
Li *et al.* [83]	In this work, colour texture retrieval method is introduced by using copula model based on Gabor wavelets. Four copula schemes are developed, and accordingly four KLDs distance of the copula schemes are introduced for colour texture image retrieval	Lack of generalization aspects, works well for texture dominated images	Two large texture databases ALOT and STex	Precision and Recall
Zhang [84]	In this work, a method combining both colour and texture features of image is proposed. Given a query, images in the database are firstly ranked using CIEL*u*v* colour space for colour features. Then the top ranked images are re-ranked according to their Gabor texture features. Results show the second process	Traditional method uses exhaustive linear search, performance is not good, moderate feature dimension	MPEG-7 for natural image retrieval test	Precision, P-R Curve

	improves retrieval performance significantly			
Höschl & Flusser [85]	Histogram-based image retrieval method which is proposed for noisy query images. Where invariant moments features in grayscale and colour histograms are taken, and finally, images are retrieved according to histogram similarity	Designed especially for additive noises, limited performance and required exhaustive search for retrieval	1,000 pictures randomly gathered from Flickr	SNR, Retrieval matching rate
Huang *et al*. [86]	In this method, the colour moment in RGB colour space in combination with the colour histogram in HSV colour space is used for colour feature extraction, the improved Zernike moments are used for shape feature extraction, and the gray level co-occurrence matrix is used for texture feature extraction. Finally fused feature are	Feature dimension is high, slow in indexing, and slow in searching time	Corel-1000	Precision, Recall, and P-R curve

	used for indexing and retrieval			

From the literature review of CBIR systems for general images presented in the Table 2.1, following observations are made:

- Most of discussed methods use exhaustive linear search in the entire database which slow the response of query.
- Most of discussed methods are slow in feature extraction and indexing and have used extra pre-processing effort
- Feature dimension is high
- Faces a problem of semantic gap between low-level features and high level understanding
- Also the CBIR performance is not so encouraging for some methods, cover a limited number of aspects, and face a problem of generalizations

2.4 Literature Review for Mammogram Classification and Retrieval

Due to the rapid growth of digital mammograms, posed a great challenge and creates a need to develop new automated tools which help radiologists to retrieve and analyse current images with past stored images [87]. There is increasing interest in the use of CBIR techniques to diagnose the stage of breast cancer, because it supports radiologists in their decision to find out similar historical cases from the pre-stored database, and help radiologist for comparing the current case with past cases [88]. Unlike text-based retrieval and manual classification, content-based approaches index images using texture, colour, and shape. Due to the complexity of the pre-processing operation, CBIR for mammogram database faces a considerable computational burden. Therefore, mammogram classification is important because it categories the images into different classes based on its appearance, and treated as a pre-processing step to speed up the accuracy, searching time and retrieval performances of CBIR systems [88-91]. Furthermore, mammogram classifications can be treated as computer-aided diagnosis (CAD) system which helps for the early diagnosis of breast cancer by detecting cancer in the mammogram. Therefore, due to the significant importance of mammogram classification for the retrieval and diagnosis, we have also gone through the state-of-art methods of this area. Table 2.2, presents the brief introduction of the state-of-art methods for mammogram classification. In this table, we briefly discussed about the used techniques, and achieved performances for the corresponding benchmark databases.

Table 2.2: State-of-art methods for mammogram classification

State-of-art methods	Used feature extraction and classification model	Accuracy & other performance measures	Used database
Srivastava *et al*. [89]	Manual cropping + CLAHE enhancement + Modified Fuzzy C-means segmentation with statistical and texture based feature extraction+ Relevant feature selection using GA-Mutual Information + SVM-MLP Classification. Also tested feature set selection using	Accuracy: 87%, Sensitivity: 95% Specificity:75%, Accuracy: 87.50%, Sensitivity: 95.83%, Specificity: 62.50%	MIAS

	sequential feature selection with k-NN + k- nearest neighbour (k-NN) Classifier		
Beura *et al.* [90]	Extracted GLCM feature from 2- level wavelet domain +Back propagation neural networks (BPNN) classifier	For MIAS: Specificity=97.0% Accuracy 94.2% For DDSM: Specificity=97.9%, Accuracy=97.4%	MIAS, DDSM
Pratiwi M *et al.* [92]	GLCM based texture feature extraction from different orientations + Radial Basis Function Neural Network (RBFNN) classifier	Highest achieved Accuracy, Sensitivity Specificity, 93.98% 94.44% 93.62% respectively	MIAS
Tzikopoulos *et al.* [93]	Combination of Fractal texture features and Statistical features, Support vector machine classifier	Accuracy=84.47-85.3%	MIAS
Buciu *et al.* [94]	Gabor wavelets and principle component analysis (PCA), Support Vector Machine classifier	Specificity=60.86%, *AUC* = 0.79, Sensitivity=97.56%,	MIAS
Prathibha *et*	DWT based texture features +	*AUC* = 0.95	MIAS

al. [95]	Nearest neighbour classifier		
Liu *et al.*[96]	Multiresolution analysis, wavelet and statistical features + Binary tree classifier	Accuracy=84.2%	MIAS
Subashini *et al.[*97]	Statistical features + Support vector machine classifier	Accuracy= 86.67%	MIAS
Wang *et al.* [98]	Histogram features+ Back propagation Neural network classifier	Accuracy=71%	MIAS
Muhimmah *et al.* [99]	Histogram features in multi resolution domain+ DAG-SVM classifier	Accuracy=77.57%	MIAS
Oliver *et al.* [100]	Textural and Morphological features + Sequential feature selection +k-NN classifier	Specificity=91%	MIAS
Miller & Astley [101]	Texture features based on Laws with Bayesian classifier	Accuracy=80%	MIAS
Karahaliou *et al.* [102]	Four different feature sets based on texture (first-order statistics, gray level co-occurrence matrices (GLCM), gray level run length matrices (GLRM) and Laws' texture energy Measures + k-NN classifier.	Best achieved classification accuracy 89%, Sensitivity 90.74% and specificity 86.96%).	DDSM

Mutaz *et al.* [103]	GLCM features + ANN Classifier	Sensitivity=91.6%, Specificity=84.17%	MIAS
Petrosian *et al.*[104]	Spatial-Gray Level Dependence and textural features + Decision tree classifier	76 % specificity, 89%-sensitivity	Local database of 180 mammograms
Kinoshita *et al.*[105]	Texture and segmented Shape features with a three layer feed-forward NN.	Accuracy=81%	Local database
Sameti *et al.*[106]	Photoelectric, Optical density, textural features with discriminant analysis classifier	Accuracy=72%	Locally collected 58 mammograms
Dhahbi *et al.* [107]	Discrete curvelet transform + t test ranking for feature selection +k-NN classifier	Accuracy=91.27% for MIAS database	MIAS and DDSM
Mudigonda *et al.* [108]	GLCM and polygon modelling (Gradient based and texture features) with jack-knife classification	Accuracy=83%	MIAS+local images (39-MIAS+15 local) total 54 images
Brijesh *et al.* [109]	Statistical features with fuzzy - neural network classifier	Accuracy=77.8–83.3%	Nijmegen database
Liyang *et al.* [110]	Statistical features in multiple view mammogram with SVM and kernel Fisher discriminant	Accuracy=85%	697 mammograms from 386

	(KFD)		cases
Szekeley *et al*. [111]	Statistical features, Texture features and combining multi-resolution Markov random models + decision trees classifier	Accuracy=88%	Local database
Alolfe *et al*. *[*112]	Forward stepwise linear regression method with a combined classifier of SVM and linear discriminate analysis(LDA)	Accuracy=82.5%	MIAS and Local database
Jona *et al*. [113]	Gray level co-occurrence matrix based statistical features + SVM classifier	Accuracy=94.0%	MIAS
G¨orgel *et al*. [114]	DWT based texture features, SVM classifier	Accuracy rate=84.8%	MIAS
G¨orgel *et al*. [115]	Stationary wavelet transform (SWT)+SVM Classifier	Accuracy rate=96.0%	MIAS
Rashed *et al*. [116]	Used fractional amount of biggest different types of Daubechies wavelets coefficients in multilevel decomposition.	Accuracy=87.06%	MIAS
Dheeba *et al*.	Pre-processing +Laws texture	Sensitivity=94.167%,	216 sample

[117]	features+ Classification using Swarm Optimized Wavelet Neural Network	Precision=92.12% Accuracy=93.67% AUC=0.9685	mammograms from 54 patients

Further, brief descriptions of some state-of-art methods for the retrieval of digital mammograms are as given:

Wei et al [118], used 4 different gray level co-occurrence matrices, constructed in order to compute each ROI in the 0°, 45°, 90°, and 135° directions, each with unit pixel distance of 5. Further, 11 Haralick statistical features are computed from each matrix and Euclidean distance based similarity measure is used for retrieval. In another contribution Wei et al [119] used Gabor filtering (Orientation=6, Frequency=4) on the underlying image, applied probability distribution and then computes Contrast, Angular second moment, Inverse difference Moment, Entropy, Variance and Correlation based statistical features to describe the textural pattern of the mammogram. Further *ED* is used as a similarity measure.

For the article Sun and Zhang [120], texture structure of mammographic image was firstly extracted by the maximum and minimum of local intensity. Then, a new texture feature based on gradient differences in 8-connected neighbours is introduced. Further combined with the weighted moments, the new descriptor was given as an index for mammographic image retrieval.

Wiesmuller and Chandy [121] has extracted gray level aura matrix from each ROI to capture the texture information. Afterwards for finding the most relevant mammograms, Euclidean distance similarity from every feature vector in the database to the query image feature vector is calculated.

Eisa et al [122] investigated the retrieval of mass and calcification mammograms from the MIAS database using texture and moment-based features.

In Quellec et al [123], the biorthogonal and orthogonal wavelets within the adaptive lifting scheme framework have been evaluated for the retrieval of mammograms from the DDSM database and is compared with the retrieval methods based on Daubechies 4-tap wavelet (Db4) and Zernike moments.

In Chores [124], classification and retrieval of benign and malignant type mammograms are given. For indexing, they used the Gabor and GLCM based texture features with addition to shape features.

Chandy et al [125] have used gray-level statistical matrix for capturing the texture features of the mammogram. Further, using this matrix they have extracted four statistical texture features for the retrieval of mammograms from mammographic image analysis society (MIAS) database. In another framework [126] they have used neighbouring feature search selection method for various combinations of texture features.

From the literature review for mammogram classification and retrieval presented in Table 2.2. and discussed subsequently, the following observation are made:

- From the literature review, it is found that most of classification and retrieval methods have used/ incorporated or introduced different variants of texture features viz. grey level statistical matrices, GLCM, SWT, Gabor filters and wavelets.
- Most CBIR methods [118-126] for mammogram are based on the manual cropping for avoiding the artifacts.
- Manually cropping of ROI from the given ground truth.

- This manual pre-processing requires too much time and is labour intensive to implement.
- Further, these retrieval methods use exhaustive linear search in the entire database. It means these approaches compare the matching (calculating the similarity measures) from all the mammograms of database, which slow down the response time.
- and have limited retrieval performances.

This thesis, addresses above mentioned issues and challenges related to design and development of CBIR system and incorporated the new approaches at pre-processing and feature extraction level, and applied the machine learning techniques, which alleviate the above issues.

2.5 Benchmark Databases

To evaluate the performance of proposed general frameworks, the experiments have been performed on *Wang* (Corel-1000, Corel-10000), and *Oliva and Torralba* (OT-Scene) image databases [127, 128]. *Wang* database contains 10 different classes of images, each class having 100 images of Africa people, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers-Roses, Horses, Hills-Mountains, and Foods-Fruits with resolution of 256*384 or 384*256 pixels. For Corel-10000 database there are 100 different class of images and each class have 100 images.

Fig. 2.1: Sample images of Wang database

Fig. 2.2: Sample images of OT database

OT scene database is created by MIT researchers which contains 2688 images of 8 different classes. These categories include: 360 images of Coast, 328 images of

forest, 260 images of Highway, 308 images of Inside city, 374 images of Mountains and hills, 410 images of Open Country, 292 images of Street , and Tall Building with 356 images. Sample images of these databases are shown in Fig.2.1 and Fig. 2.2.

- *Mammographic image analysis and society (MIAS) database* [129]

To evaluate the performance of retrieval frameworks for the diagnosis of breast cancer using mammogram, mammographic images analysis society (MIAS) is taken. In MIAS database ground truth details of mammographic images (size 1024*1024 pixels) are given with corresponding ids. This database contains 115 images from Abnormal and 207 images from Normal classes. During retrieval reported ground truth information are used for the identification of normal and abnormal masses. Further abnormal mammogram is classified into 6 different categories, including 19 images in architectural distortion, 15 in asymmetry, 29 images in calcification, 15 in circumscribes masses, 15 in ill-defined masses and 19 in speculated mass.

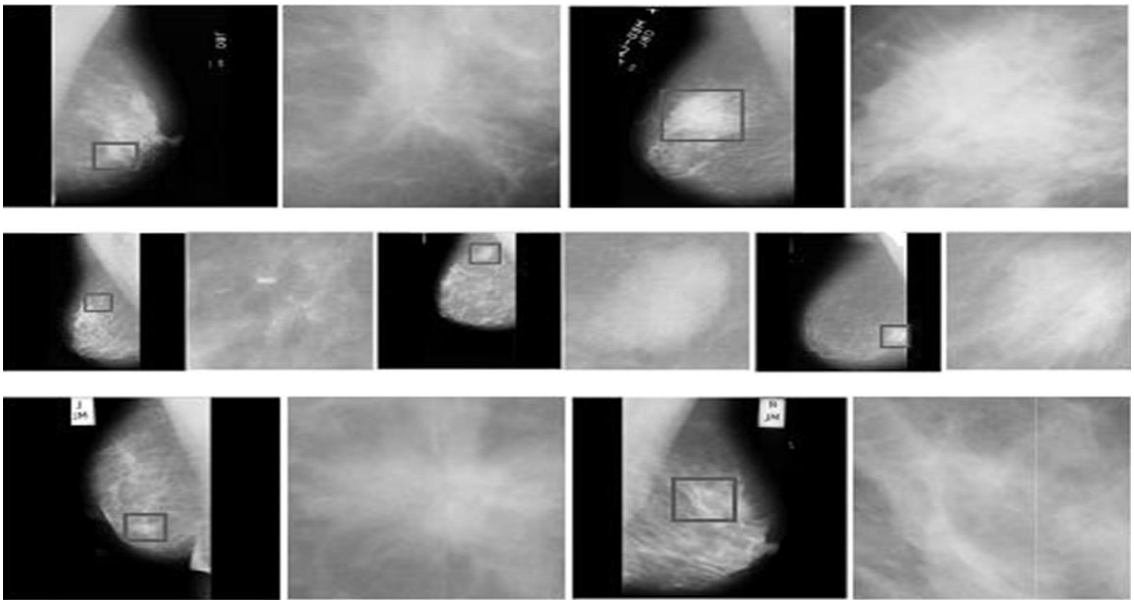

Fig. 2.3: Sample mammograms of different classes

Fig. 2.3 shows the sample mammograms and their manual cropped ROIs from the MIAS database containing the following classes. From left to right and top to bottom: architectural distortion (ARCH), asymmetry (ASYM), calcification (CALC), circumscribed masses (CIRC), ill-defined masses (MISC), speculated masses (SPIC) and normal [125].

2.6 Performance Metrics

Confusion Matrix: The *confusion matrix* is a useful tool for analysing how well your classifier can recognize images of different classes:

Table 2.3: Confusion matrix to measure the performance of the classifiers

Predicted Class			
	Positive	Negative	Total

Positive	TP	FN	P
Negative	FP	TN	N

The performance of the classifiers is measured by the quantity of True positive (*TP*), True Negative (*TN*), False Positive (*FP*), False Negative (*FN*). Where *TP* (True Positive) is the number of positive instances that are classified as positive, *FP* (False Positive) is the number of negative instances that are classified as positive, *TN* (True Negative) is the number of negative instances that are classified as negative and *FN* (False Negative) is the number of positive instances that are classified as negative. By using these quantities standard accuracy, sensitivity, specificity, precision, *MCC* and *ROC* area performance measures are defined as:

Precision: Precision is defined as the proportion of instances classified as positive that are really positive.

$$\text{Precision } = \frac{\text{TP}}{\text{TP+FP}} \quad (2.9)$$

For retrieval context this precision can also be defined as:

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of retrievd images}} \quad (2.10)$$

Recall: Recall is defined as the proportion of positive instances that are correctly classified as positive.

$$\text{Recall (Classifier)} = \frac{\text{TP}}{\text{TP+FN}} \quad (2.11)$$

This recall can be defined as:

$$\text{Recall(CBIR)} = \frac{\text{Number of relevant images retrieved}}{\text{Number of relevant image in database}} \quad (2.12)$$

Accuracy: Accuracy is defined as the proportion of instances that are correctly classified.

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \tag{2.13}$$

FP **Rate:** It is the probability that a classifier produces erroneous results as positive results for negative instances.

$$FP\ Rate = \frac{FP}{FP+TN} \tag{2.14}$$

F-Measure: This measure is approximately the average of the two when they are close. The two measures are sometimes used together in the F1 **Score** (or **f-measure**) to provide a single measurement for a system.

$$F - Measure = \frac{2\ (Precision \times Recall)}{(Precision + Recall)} \tag{2.15}$$

Matthew's correlation coefficient (*MCC*): The *MCC* is a balanced measure that considers both true and false positives and negatives. The *MCC* can be obtained as

$$MCC = \frac{(TP)(TN)-(FP)(FN)}{\sqrt{[TP+FP][TP+FN][TN+FP][TN+FN]}} \tag{2.16}$$

Receiver operating characteristics (*ROC*): The *ROC* is a graph that shows the performance of a classifier by plotting *TP* rate versus FP rate at various threshold settings. Area under *ROC* curve (*AUC*) of a classifier is the probability that the classifier ranks a randomly chosen positive instance higher than a randomly chosen negative instance.

Speedup: Proposed work, searching time improvement as compare to conventional exhaustive search can be estimated using Speedup metric, which are calculated using following formula:

$$Speedup = \frac{Total\ number\ of\ images\ in\ the\ database}{Total\ Number\ of\ images\ in\ the\ cluster} \tag{2.17}$$

Random Index (RI): The random index measure is used for the evaluation of segmentation and clustering algorithms. The random index between ground truth image (GT) and segmented image (S) is calculated by summing the number of pixel pairs with same label and number of pixel pairs having different labels in both S and GT, and then dividing it by total number of pixel pairs. RI values lie between 0-1, where higher value of RI indicates perfect segmentation.

2.7 Conclusions

In this chapter, the theoretical backgrounds related to content-based image retrieval as well as literature review are presented. At first, brief overview of feature extraction methods, similarity measures were presented, which provided basis for computer vision applications as discussed in subsequent chapters of the thesis. Further, in this chapter a literature survey of prominent approaches for of general CBIR system as well as for mammogram application was discussed and research gaps were identified.