A RESEARCH ON STATISTICAL TEXTURE ANALYSIS METHODS AND THEIR APPLICATION

A MONOGRAPH
BY

Dr. SREEJA MOLE S.S.
Professor and Head,
Department of Electronics & Communication Engineering,
Christu Jyothi Institute of Technology and Science, Jangaon, India.
Email: sreebommy@gmail.com

IJCESR
2394-0697

Publishing Partner:
TROINDIA

TABLE OF CONTENTS

ABSTRACT

Texture analysis is moving towards further technical advancements in both logically and accurately in the digital image processing. In that, image classification is the critical step in image analysis and pattern recognition, which is one of the most difficult tasks in image processing to determine the quality of the final result of analysis. The main aim of texture analysis is to improve the tonal information of the image, since texture is a major property of an image that represents important information about the structural arrangement of features in an image. Because texture determines the overall visual smoothness of the image features and the rate of texture is to create a brightness value for each pixel in an image. Texture is measured statistically using a moving window through out the image. In this study, the different methods of statistical texture analysis have been analyzed and the comparative analysis also has been illustrated.

The existing results of texture algorithms show the benefit of texture in supervised and unsupervised classification. Statistical operators including skew ness, kurtosis, variance, standard deviation, maximum and mean, Euclidean distances are used for texture analysis image processing to characterize the distribution of gray levels in an image. A number of various texture analysis methods have been introduced namely statistical, structural, transform based and model based method. Normally textures are studied through statistically and syntactically method. The statistical method measures the coarseness and directionality of texture in terms of averages on a window of the image. But syntactical method describes the shape and distribution of the entities. For achieving better classification accuracy, it is inevitable to study the main features of the statistical method that includes contrast, entropy, homogeneity, auto correlation function, power spectrum, different gray level statistics and co-occurrence matrices. Texture analysis also plays an important role in many image analysis applications. In industrial visual inspection, texture information can be used in enhancing the accuracy of color measurements. Texture methods can also be used in medical image analysis, biometric identification, remote sensing, content based image retrieval, document analysis, environment modeling, texture synthesis and model based image coding. This monograph makes a number of contributions. The performances of the proposed classification methods have been compared with other existing methods.

In the first study a new statistical method of texture analysis has been presented which is focused on texture characterization and discrimination with features like local binary pattern operator, texture spectrum operator and entropy based local descriptor. In order to evaluate the performance of the texture spectrum in texture characterization and classification the experimental study have been carried out using Brodatz natural images where the local texture information for a given pixel and the neighborhood is characterized by the corresponding texture unit. The global textural aspects of an image are revealed by its texture spectrum. Here supervised classification method has been used to extract the textural information of an image with respect to texture characteristics. The designed work has been compared with other existing methods for performance analysis.

The second study of this monograph presents a new classification technique named unsupervised hybrid classification for texture analysis (UHCTA) that integrates different unsupervised methods. This method has the properties which are flexible to achieve higher classification rate. The comparative analysis also has been done for both fixed and varying window sizes. When using co-occurrence matrices in discriminating different textures the classification accuracy increases to a satisfactory level but with the high computational cost. Moreover to improve the classification accuracy it is a must to increase the window size in addition with the selection of optimal window size.

When using statistical approach in texture analysis for image classification more complication is to be met. Particularly gray level co-occurrence matrix approach is applied in discriminating different texture images that compromises the classification rate. The third study of this monograph presents an image classification method based on pixel by pixel with maximum likelihood estimates that must be compared to a single window classification not only to monochrome images but with the color images too. The idea of proposed method is to consider the relative frequencies for which two neighboring pixels are separated by a distance on the image.

The fourth study of this monograph presents the both supervised and unsupervised statistical method for the detection of micro-calcification on mammogram. Texture analysis has been very much used in medical image problems. Among all medical image tasks, detecting the micro-calcification on mammogram is the most difficult one. Moreover micro-calcification are deposits of calcium that can be seen in mammogram and owing to the small size of micro-calcification with a diameter of less than 0.5mm level and are in the form of groups as clusters in homogeneous background it is very difficult to detect. The common set of mammogram images has been taken from MIAS to identify the classification accuracy.

On the whole, the texture analysis methods designed in this dissertation are very useful in providing higher classification accuracy. The results suggest that the proposed classification methods for texture analysis perform significantly better in the way of incorporating remarkable accuracy measures with reliability. The approach would definitely increase its usability and adds more to the future prospects of improving classification in texture analysis for many applications.

CHAPTER 1
STATISTICAL TEXTURE ANALYSIS AND THEIR APPLICATION

1.1 INTRODUCTION

Texture classification is an important research area in digital image processing. The main aim of texture analysis is to improve the tonal information of an image, since texture is a major property of an image that represents important information about the structural arrangement of features in an image. Texture is measured statistically using a moving window through out the image. Statistical operators including skewness, kurtosis, variances, standard deviation, maximum and mean, Euclidean distances are used for texture analysis. A number of various texture analysis methods have been introduced namely statistical, structural, transform based and model based method. The statistical method measures the coarseness and directionality of texture in terms of averages on a window of the image. But syntactical method describes the shape and distribution of the entities. For achieving better classification accuracy, it is inevitable to study the main features of the statistical method that includes contrast, entropy, homogeneity, auto correlation function, power spectrum, different gray level statistics and co-occurrence matrices. Texture analysis plays an important role in many image analysis applications. In industrial visual inspection, texture information can be used in enhancing the accuracy of measurements. Texture methods can also be used in medical image analysis, biometric identification, remote sensing, content based image retrieval, document analysis, texture synthesis and model based image coding.

1.2 NEED FOR THE PRESENT STUDY

Texture is an important spatial feature useful for identifying objects. Texture, an intrinsic property of object surface, is used by the visual perception system to understand a scene; therefore texture analysis is an important component of image processing. But texture is not easy to quantify, since there is neither a consensus on its definition nor a precise mathematical formulation for it. When the visual examination of an image suggests that the basis for discriminating various structural regions is a texture rather than color or brightness, it may be possible to use simple texture operators. The goal of texture classification is to discriminate different classes of textures. Textures cannot be defined at single pixel level but always associated with image regions. Texture can be evaluated as being fine, coarse or smooth rippled, irregular of limited. Texture is an innate property of virtually all surfaces. The grain of wood, the weave of a fabric, the pattern of crops in a field etc and it contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment. The textural properties of images appear to carry useful information for discrimination purposes it is important to develop features for textures. Major approaches for texture classification are based on a discrimination function using several texture characteristics. The first task is to extract texture features which most completely embody information of texture in the original image. Several methods exist such as extracting the feature either directly from the image statistics or from the spatial/frequency domain. The texture analysis problems include classification and segmentation of texture regions in a scene.

In statistical texture analysis, supervised and unsupervised method are available and have advantages and disadvantages. In the unsupervised approach the Texture Spectrum Operator

(TSO) has the advantage that texture aspects of the images are characterized by the corresponding texture spectrum instead of a set of texture measures. On the other hand Local Binary Pattern (LBP) operator is attractive but the implementation difficulty in the form of delta value definition form a user to set a threshold value, which makes it dependent on the gray scale values. Gray Level Co-occurrence Matrix (GLCM) method is even though suitable for unsupervised texture classification analysis but with the limitation by classifying large primitives. Finally Entropy Based Local Descriptor (EBLD) method has the disadvantage of considering the data in terms of probability resulting uncertainty and instability.

So it is important to study the texture analysis of both supervised and unsupervised methods in order to obtain a high degree of classification accuracy. Since texture analysis plays an important role in all fields effectively and particularly in medical applications and so it is inevitable to study the utilization of different classification methods and to compare.

1.3 OBJECTIVES OF THE PRESENT STUDY

Texture classification is a fundamental problem in computer vision with a wide variety of application, how to characterize textures using derived feature, how to define a robust distance, similarity between textures. Textural features must be invariant to image variation and at the same time sensitive to intrinsic spatial structures that define textures. Because there is no obvious feature common to all texture images, texture features are often proposed based on assumptions for mathematical convenience.

The following are the high level objectives of the present study.

- To propose suitable descriptors for texture, based on LBP and Entropy.
- To perform unsupervised classification by clustering pixels in a data set, based on statistics only without any user-defined training classes.
- To perform supervised classification using the proposed descriptors and compare the results with existing operators.
- Suggesting methods to improve the classification accuracy by means of integrating the features.
- Applying the statistical texture quantification methods on medical images to determine the suspicious masses or deposits particularly in mammograms.
- Comparing the different classification methods on accuracies.

1.4 CLASSIFICATION OF TEXTURE IMAGES

Texture classification is a key field in many computer vision applications ranging from quality control to remote sensing. The problem of feature selection is defined by a given a set of candidate features; select a subset that performs the best under classification system. This procedures can reduce not only the cost of recognition by reducing the number of features that need to be collected, but in some cases it can also provide a better classification accuracy due to finite sample size effects. A textured analysis task generally involves classification. Here, the generated features span a high dimensional feature space which is sub divided into a set of classes. A feature vector is then assigned a class label according to its position in feature space. When labeled data is available a priori, the classification is called supervised otherwise it is called unsupervised. An important task is supervised texture image classification. Here, each image is represented by a single feature vector. A training set containing several labeled images. By applying classification on this set, texture classes are defined according to the labels of the textures in the training set

Texture classification refers to the process of grouping test samples of texture into classes where each resulting class contains similar samples according to some similarity criterion.

1.4.1 Supervised Classification

The goal of classification in general is to select the most appropriate category for an unknown object. The Supervised Classification is based on some discrimination function using several image features. First the data samples are collected and properly normalized from all classes. Supervised classification procedures are the essential analytical tools used for the extraction of quantitative information. A supervised classification needs existing reference data to establish user-selected training sets.

1.4.2 Unsupervised Classification

The second major approach to image classification is the unsupervised clustering of image data, where no a priori information is required. A large number of clustering algorithms are available in the literature. Cluster analysis is viewed as a process of partitioning an image into groups such that patterns belonging to the same group are similar to each other than the patterns belonging to different groups. Unsupervised classification is a means by which pixels in an image are assigned to spectral classes without the user having knowledge of the existence or names of those classes. Clustering analysis is usually the common method to perform an unsupervised image classification. The user determines the number and the location of the spectral classes into which the data falls, as well as the spectral class of each pixel according to statistical similarity. After the completion of the clustering, it is the task of the analyst to identify those classes by associating a sample of pixels in each class with available reference data. Often unsupervised classification is used on its own, particularly when the training data for supervised classification cannot be obtained or is too expensive to acquire.

An essential problem in computer vision is texture analysis and they are characterized into two types, namely, structural and statistics. Many texture classification methods already have been implemented and each one has its own advantages and disadvantages. Firstly the TSO method has the advantage that the texture aspects of an image are characterized by the corresponding texture spectrum instead of a set of texture measures and the texture spectrum can be directly used for image classification. On the other hand LBP method is attractive to some extent but with the implementation difficulty in the form of delta values definition, form a user to set the threshold values which makes it dependent on the gray scale values. The GLCM method is even though suitable for unsupervised texture classification analysis but with the limitation in classifying the large primitives. The EBLD method has the disadvantage of considering the data in terms of probability resulting uncertainty and instability of data.

1.5 APPLICATIONS OF TEXTURE ANALYSIS METHODS

Texture analysis methods have been utilized in a variety of application domains. In some of the mature domains (such as remote sensing) texture already has played a major role, while in other disciplines (such as surface inspection) new applications of texture are being found. The main areas of applications include the texture in automated inspection, medical image processing, and document processing and remote sensing. The role that texture plays varies depending upon the application. Some of the texture analysis applications are described here.

Inspection

There have been a number of applications of texture processing to automated inspection problems. These applications include defect detection in images of textiles and automated inspection of carpet wear and automobile paints. In the detection of defects in texture images, most applications have been in the domain of textile inspection. The features they use to perform classification is based on tonal features such as mean, variance, skewness, and kurtosis of gray levels along with texture features computed from gray level co-occurrence matrices. The combination of using tonal features along with textural features improves the correct classification rates over using either type of feature alone.

Medical Image Analysis

Image analysis techniques have played an important role in several medical applications. In general, the applications involve the automatic extraction of features from the image which is then used for a variety of classification tasks, such as distinguishing normal tissue from abnormal tissue. Depending upon the particular classification task, the extracted features capture morphological properties, color properties, or certain textural properties of the image. The textural properties computed are closely related to the application domain to be used. Some diseases, such as breast cancer in such a manner that the resulting changes in the X-ray images are texture changes as opposed to clearly delineated lesions. In such applications, texture analysis methods are ideally suited for these images.

Document Processing

One of the useful applications of machine vision and image analysis has been in the area of document image analysis and character recognition. Document processing has applications ranging from postal address recognition to analysis and interpretation of maps. In many postal document processing applications (such as the recognition of destination address and zip code information on envelopes), the first step is the ability to separate the regions in the image, which contains useful information from the background.

Remote Sensing

Texture analysis has been extensively used to classify remotely sensed images. Land use classification where homogeneous regions with different types of terrains (such as wheat, bodies of water, urban regions, etc.) need to be identified is an important application.

1.6 PROPOSALS AND ORGANIZATION OF THE MONOGRAPH

In this work, the different texture analysis methods to improve the classification accuracy using supervised and unsupervised methods and their application are studied .The following modules have been formulated.

- Studying the supervised classification of images using textural features.
- Improving classification accuracy using unsupervised hybrid methods.
- Studying the classification on color images of statistical texture images using pixel by pixel based mechanism.
- Comparative analysis of different classification methods in texture analysis.
- Detecting the Micro-calcification on Mammogram using supervised and unsupervised methods.

The literature survey pertaining to this monograph is given in **Chapter 2**.

In the **Chapter 3** the local texture information for a given pixel and the neighborhood is characterized by the corresponding texture unit. The global textural aspects of an image are revealed by its texture spectrum. Here supervised classification method has been used to extract the textural information of an image with respect to texture characteristics. The purpose of this phase is to present a new statistical method of texture analysis which is focused on texture characterization and discrimination with features like LBP, TSO and EBLD. In order to evaluate the performance of the texture spectrum in texture characterization and classification, the experimental studies have been carried out using Brodatz natural images for both supervised and unsupervised classification methods.

When using co-occurrence matrices in discriminating different textures the classification accuracy increases to a satisfactory level but with the high computational cost. Moreover to improve the classification accuracy it is a must to increase the window size in addition with the selection of optimal window size. In the **Chapter 4**, the methods only used a single window classification for each pixels resulting unsatisfactory classification accuracy. So in the second phase a new classification technique named Unsupervised Hybrid Classification for Texture Analysis (UHCTA) was used that integrates different unsupervised method. This method has the properties, which are flexible to achieve higher classification rate. Different images of Brodatz data set have been taken for the comparative analysis for both fixed and varying window sizes.

When using statistical approach in texture analysis for image classification more complication is to be met. Particularly grey level co-occurrence matrix approach is applied in discriminating different texture images that compromises the classification rate. So in the **Chapter 5** of study, it is necessary to concentrate on image classification method based on pixel by pixel with maximum likelihood estimates that must be compared to a single window classification not only to monochrome images but with the color images too. The idea of proposed method is to consider the relative frequencies for which two neighboring pixels are separated by a distance on the image. Texture measures such as homogeneity, contrast and entropy are derived from the co-occurrence matrix. The color image is represented by three co-ordinates namely luminance and two chrominance.

Texture analysis has been very much used in medical image problems. Among all medical image task in detecting the micro-calcification on mammogram is the most difficult one. Since breast cancer is the most prevalent cancer that leads to death of woman today. Moreover micro-calcification are deposits of calcium that can be seen in mammogram and owing to the small size of micro-calcification with a diameter of less than 0.5mm level and are in the form of groups as clusters in homogeneous background it is very difficult to detect. So in **Chapter 6,** both supervised and unsupervised statistical methods have been applied on mammogram images for the detection of micro-calcification. The common sets of mammogram images have been taken from MIAS and experiment with the supervised and unsupervised methods to determine and study the classification accuracy.

Chapter 7 sums up the conclusion and discusses future enhancement of this research work.

CHAPTER 2
LITERATURE SURVEY

In this chapter, the vast literature contributing to the improvement of methods for classification accuracy in texture analysis with supervised and unsupervised methods are presented. The literature on applications on texture analysis was also reviewed. Subsequently, a brief review has been made on case studies relative to the different texture, color and pixel by pixel methodologies to improve the classification rate.

Image texture, defined as a function of the spatial variation in pixel intensities (gray values), is useful in a variety of applications and has been a subject of intense study by many researchers. One immediate application of image texture is the recognition of image regions using texture properties. Texture is the most important visual cue in identifying these types of homogeneous regions and the goal of texture classification is then to produce a classification map of the input image where each uniform textured region is identified with the texture class it belongs to. The second type of problem that texture analysis research attempts to solve texture segmentation in order to obtain the boundary map. Texture synthesis is often used for image compression applications. It is also important in computer graphics, where the goal is to render object surfaces which are as realistic looking as possible. The shape from texture problem is one instance of a general class of vision problems.

2.1 BASICS ON TEXTURE ANALYSIS

Haralick et al (1973) described a class of quickly computable textural features which are based on statistics. They also used in category identification tasks of three different kinds of image data. Their results shows that the easily computable texture features have a general applicability of image classification applications. Even though the methods are suitable for variety of image data but with the necessity to determine the size of the sub-image regions and the distance.

Richard Coners and Charles Harlow (1980) evaluated the four texture analysis algorithm to perform automatic texture discrimination are Spatial Gray Level Dependence Method(SGLDM),Gray Level Run Length Method(GLRLM),Gray Level Difference Method(GLDM) and the Power Spectral Method(PSM). Of the four algorithms, it is discovered that the SGLDM is the most powerful algorithm and GLDM is also more powerful than PSM. Since SGLDM uses the class of Markov textures from 180^{o} of rotation of texture.

Dong Chen He and Li Wang (1990) proposed to demonstrate the usefulness of the texture spectrum for texture classification. The results have been very attractive when applying the texture spectrum to classify Brodatz natural images. From their evaluation, it is identified that the texture spectrum is sensitive to the directional aspects of texture.

Dong Chen He and Li Wang (1991) presented textural filters based on the textural spectrum. Since Conventional Digital Filtering Techniques based on classical Fourier analysis are widely used in digital image processing and unsatisfaction may be encounted when applying filters to texture analysis of images with the requirement of some specific filters. The authors applied the Brodatz natural images in order to show a promising potential of the textural spectrum for the design of the textural filters.

Dong Chen He and Li Wang (1991) proposed the texture spectrum for a unsupervised classification of texture images. Since it is also a good way to test the discriminate performance features used in unsupervised classification. The authors also introduced the texture spectrum to the natural classification of images with the promising results.

Stachowicz and Lemke (2002) proposed a system for image segmentation and classification using color as the primary feature. This system is comprised of two phases namely segmentation and classification. In the first step, an image is searched with a detection algorithm to determine the location of any possible foreground elements. These areas are extracted from the image to be used in the next step. Classification is done using a set of eight color features that are optimally selected for each database. The appropriate feature vector is created for each foreground area removed from the original image. The correlated vector is then compared to a reconstructed database to be identified.

Veronica (2007) explained the five different data clustering algorithms namely K-Means, Fuzzy C Means, Mountain subtractive and Partition Simulation FCM (PsFCM). From their studies, the performance of K-means algorithm depends on the position of initial cluster centre. The FCM algorithm depends on initial cluster centers. The PsFCM algorithm is an extension of FCM that reduces the number of iterations required by selecting the initial cluster center that is very close to the actual cluster centre of the data set. The mountain clustering method is an effective approach to approximate the estimation of cluster centers on the basis of density measure. The subtractive clustering eliminates the problem of mountain clustering, where its computation grows exponentially with the dimension of the pattern over all grid points.

Xiuwen Li and Deliang (2003) demonstrated that the spectral histogram provides sufficient features statistic for texture classification. For that they proposed a filter section algorithm for texture classification. By obtaining satisfactory classification results on natural texture data sets. From the comparison, it is found that the spectral histogram improves the classification performance significantly.

Cun Lu Xu and Yan Qiu Chen (2004) proposed the use of information derived from the graph of a texture image function for texture description. The proposed Statistical Landscape Features (SLF) carries five frequently used techniques namely the autocorrelation, the edge frequency measures, the Spatial Gray Level Dependence Matrix (SGLDM), the Statistical Geometric Feature (SGF) and the use of the Discrete Wavelet Transform (DWT). Among the above the SLF achieves a very high correct classification rate of 94.7% on the entire Brodatz data set. Besides the very high performance, another remarkable advantage of the proposed method is that it has no parameters to tune.

Srinivasa and Thomas et al (2005) presented a Knowledge Discovery and Data mining (KDD) process that includes preprocessing, transformation, data mining and knowledge extraction. The two important tasks of data mining are clustering and classification and are done with a proposed generic feature extraction for classification using Fuzzy C-means (FCM) clustering. The raw data is preprocessed and then data points are clustered using the fuzzy C-means technique. Feature vectors for all the classes are generated by extracting the most relevant features from the corresponding clusters and used for further classification. Artificial neural network and support vector machines are used to perform the classification task. Experiments are conducted on four datasets and the accuracy obtained by performing specific feature extraction for a particular data set is compared with the generic feature extraction scheme. The algorithm performs relatively well with respect to classification results, when compared with the specific feature extraction technique.

Chin Cheng Hung et al (2006) aimed at analyzing and comparing some of the simple but powerful spatial image classification to explore the strengths and weakness in remote sensing applications. The comparison of texture spectrum and local binary pattern also was done. The authors also used co occurrence probabilities to increase the performance by increasing the window size in order to achieve good amount of classification.

Fok Hing Chi Tivive and Abdesselam Bouzerdoum (2007) presented an architecture that can be used as a nonlinear feature extractor for texture segmentation which comprises two layers of feature extraction units. Each layer is arranged into several planes called feature maps and the features extracted from the second layers are used as the final features. Combining the nonlinear features extracted with a classifier, the authors developed a texture segmentation system that does not depend on predefined filters for feature extraction. Tested on the Brodatz texture images, the proposed texture segmentation systems achieve better classification accuracy.

Umarani et al (2007) explained the combined approach for unsupervised texture classification of statistical and structural methods by conducting Nair's test. The spectrums were shown to be effectively used for unsupervised texture classification for Brodatz and Meastex data base of texture image with a correct classification of 96%. They also found the reason for misclassification which is due to the overlapping the regions and the influence of the tolerance part.

Bor-Chen Kuo et al (2008) proposed a new fuzzy clustering, namely Fuzzy C-Weighted Mean (FCWM). The cost function of the classical Fuzzy C-Mean (FCM) is defined by the distances from data to the cluster centers with their fuzzy memberships. From Nonparametric Weighted Feature Extraction (NWFE) was introduced to establish a novel FCM-like clustering algorithm in their study. The real data experimental results show that the proposing FCWM outperforms the original FCM.

Srinavasan and Shobha (2008) presented an overview of the methodology of algorithm for statistically texture analysis. Different methods for digital image texture analysis have also been reviewed including first order statistics based approach and frequency based texture analysis. From their work, it is found that spectral techniques are ideally suited for describing the directionality of periodic or almost periodic 2D patterns in an image.

Alaa Eleyan Hasan Demire (2009) introduced a new face recognition method based on the Gray Level Co occurrence Matrix method (GLCM). Two methods have been used to extracts feature vectors from the GLCM for face classification. The first method extracts the well known Haralick's features to form the feature vector, where the second method directly uses GLCM by converting the matrix into vector. From their results, it is come to know that the proposed GLCM based face recognition system outperforms the other methods namely principal component analysis and linear discriminant analysis.

Toure Land Zou Beijimusau (2010) brought out a method for segmentation of color images based on fuzzy classification. It proceeds in a first step by a fine segmentation using the algorithm of FCM. The method then applies a test fusion of fuzzy classes. The result is a coarse segmentation, where each region is the union of elementary regions grown from FCM. The FCM clustering is an iterative partitioning method that produces optimal c-partitions and the standard FCM algorithm takes a long time to partition a large data set. The proposed FCM program must read the entire data set into a memory for processing. Their results show that the system performance is robust to different types of images.

2.2 SUPERVISED CLASSIFICATION

In the previous section, the basics on texture analysis, features have been discussed. This section deals with the supervised classification methods handled by different authors in order to improve the classification accuracy.

Haralick (1979) surveyed an image processing literature on the various approaches and models for textures. They concluded that the statistical seems to work well for micro- textures. The author also concluded that the textures seem to work well using histograms of primitive properties.

He and Li Wang (1990) presented a new statistical method of texture analysis based on the concept of texture unit and texture spectrum. They evaluated that the texture spectrum is able to reveal texture information in digital images. The authors also explained the advantages of statistical method of texture analysis.

Argenti et al (1990) presented a fast algorithm for calculating parameters of co-occurrence matrices. This method has been applied to the problem of classification and segmentation of artificial and natural scenes. For the classification co-occurrence matrix parameters have been taken to implement pixel by pixel using supervised learning and maximum likelihood estimates. The problem of texture boundary recognition has also been considered. Experimental results show that the improvement of classification rate can be achieved by using co-occurrence matrices method, when compared to a single window classification.

Nicolas Vanden Broucke (2004) described a new approach to color image segmentation, which is considered as a supervised pixel classification problem. The authors proposed an original color image segmentation method based on pixels. They also presented a supervised learning scheme to determine the most discriminative color texture features and described the texture classification.

Timo Ojala and Matti Pietikainen (2002) discussed and presented the effective multi-resolution approach to classify the texture based on local binary pattern. The authors termed the local binary patterns as uniform, which are the fundamental properties of local image texture. From their experimental results, it is found that these operators characterize the spatial configuration of local image textures with better performance.

Chunmei Liu et al (2005) proposed an algorithm for detecting texts in images and video frames. It is performed by three steps: edge detection, text candidate detection and text refinement detection. Firstly, it applies edge detection to get four edge maps in horizontal, vertical, up-right, and up-left direction. Secondly, the feature is extracted from four edge maps to represent the texture property of text. Then k-means algorithm is applied to detect the initial text candidates. Finally, the text areas are identified by the empirical rules analysis and refined through project profile analysis. Experimental results demonstrate that the proposed approach could efficiently be used as an automatic text detection system, which is robust for font size, font color, background complexity and language.

Shu Liao and Albert (2007) proposed a new feature extraction method, which is robust against rotation and histogram equalization for texture classification. The proposed method has three novel contribution a) Advanced Local Binary Pattern (ALBP) approach captures the most essential local structure characteristics of texture images. b) It extracts global information by using matrix measure based on the spatial distribution information of the dominant pattern introduced by the ALBP. c) It is robust to histogram equalization. The authors also compared with other widely used texture classification techniques and evaluated by applying classification test.

Gobert Lee and Mioshi Fujita (2007) investigated the use of K Means clustering an unsupervised classifier, which does not depend on the label of data for classification. K Means clustering is an unsupervised classifier, which eliminates the dependency of class labels in the data but may be subjected to loss of information, which could affect the goodness of the classifier. The proposed ROC generating procedure enables K-Means clustering to be evaluated against other classifier of the common ground.

Albergne Ella et al (2008) performed an experiment to compare the performance of the Gray Level Co occurrence Matrix (GLCM) with that of other texture features. The performance of the texture feature is measured by computing the classification accuracy achieved on a supervised set of images spread over eight settlement classes focusing on informal low cost housing. The objective of their research is to establish which of the well texture feature algorithm are most suitable for automatic classification. From their results, it is identified that GLCM performs very well and Local Binary Pattern texture features have a small advantage in terms of classifying images.

Hui Zho et al (2008) proposed a novel extended LBP operator for texture analysis. The new LBP operator classifies and combines the local pattern based on analyzing their structure and occurrence probabilities. They used the Brodatz texture database to show the performance improvement for the novel extended LBP operator. The experimental results demonstrate that the LBP is more robust against noise than original LBP operator and it has better discrimination ability as well as better classification accuracy.

Shu Liao and Albert (2007) proposed a novel approach for the dominant local binary pattern as a texture classification. They found the solution even though in the toughest situation (SNR=5), the proposed method effectively extract the existing texture information to produce promising classification rates. In their work the classification performance has been demonstrated in three data bases which contain low image resolution textures, similar appearance textures and texture set with large number of classes.

Jaime Melendez and Domenecpuig (2010) presented a new efficient technique for supervised based texture classification. The proposed scheme first performs the selection process that automatically determines a subset of prototype that characterizes each texture based on the outcome of a multichannel Gabor multifilter bank. Then every image feature is classified into one of the given texture classes by using a K-NN classifier fed with the prototype and is compared to previous texture classifier by using both Brodatz and real outdoor texture images. From their work, it is noted that over filters and normalized cut clustering which achieves good classification result with low computation time.

Salahuddin et al (2009) proposed a new method for texture classification which is extremely fast due to the low dimensionality of feature space to extract distinctive features at a very early stage. To remove, Surface texture classification is an important aspect of computer vision and a well studied problem. From their experimental results it is known that the increase in speed was achieved for texture classification while maintaining accuracy.

Bhowmick and Chattopadhyay (2009) presented a content based video indexing and retrieval traces back to the elementary video structures, such as a table of contents. Thus, algorithms for video partitioning have become crucial with the unremitting growth in the prevalent digital video technology. This demands for a tool which would break down the video into smaller and manageable units called shots. The authors proposed a shot boundary detection technique for abrupt scene cuts. This method computes co-occurrence matrices by taking block differences between the consecutive frames in each of R, G, and B plane, using Sum of Absolute Differences (SAD).

2.3 UNSUPERVISED CLASSIFICATION

Previous section dealt with the literature survey on different types of supervised classification methods. This section focuses attention on unsupervised classification methods.

Andreas Teuner et al (1995) proposed a novel method for efficient image analysis that uses tuned matched Gabor Filters. The algorithmic determination of the parameters of Gabor filters is based on the analysis of spectral features. Contrast obtained from interactive computation of pyramidal Gabor transforms with the help of unsupervised classification method has been more attractive.

Timo Ojala and Pietikainan (1999) proposed a solution to unsupervised texture segmentation which uses distribution of local binary pattern and pattern contrast to measure the similarity of adjacent image regions during the segmentation process. The authors also evaluated the performance of the methods with various types of test images. The proposed method can be easily generalized to utilize other texture features, multiscale information and color features.

Alvin and William (2000) proposed a general multi-scale approach towards image segmentation. Clusters in the joint spatial feature domain are assumed to be the properties of underlying classes and the recovery of which is achieved by the use of mean shift procedure. The complete algorithm is applied to perform color and texture segmentation on both synthetic and real images. The algorithm is also flexible due to its ability to control segmentation sensitivity and robust through the use of the mean shift procedure and multiscale procedure. From the results, it is understood that the algorithm produces segmentation with smooth regions of homogeneity behavior and accurate boundaries which are especially useful in the higher level task of object recognition and scene analysis.

Padma Priya et al (2002) proposed a novel technique to implement an algorithm for unsupervised texture segmentation. LBP has been demonstrated to provide a robust and efficient framework for texture segmentation. The proposed algorithm uses a new data structure. The algorithm has been tested using mosaic images consisting of various textures in addition to natural scenes. From the results of the geometrical natural images, it is evident that the color has an important contribution to the discriminative power of the features.

Nicolas et al (2004) described a new approach for color texture feature extraction and selection. The color texture feature is computed by taking window account the color components of the pixels. The authors determine the most discriminating color texture features among a multidimensional set of color by means of interactive feature selection procedure. The soccer image segmentation is achieved by pixel classification and the algorithm includes the color texture features which are processed in the neighborhood of the pixels through the unsupervised approach.

Subramanian Rallaband petal (2005) derived an unsupervised classification algorithm by modeling observed data in a mixture of several mutually classes that are described by linear combination of independent non-Gaussian densities. In their algorithm, parametric non linear functions were used to improve classification accuracy compared with standard Gaussian mixture models. It is also concluded that the same approach can be used to identify multiple classes in a single image.

Matti Pietikainen (2005) studied the texture analysis problem and proposed a new method for view based recognition of 3D texture surfaces using multiple local pattern histograms as texture models. The author also considered both shape and texture information to represent face analysis. From their results, it is concluded that the texture and the ideas behind the local binary pattern methodology would have much vital role in image analysis and computer vision.

Yi Li et al (2006) proposed a new method based on permutation entropy and gray level features. Permutation entropy is a new complexity measure for time series based on comparison of neighboring values. In their work, the grey level features are introduced with the combination of four dimensions. Permutation entropy was used to construct 6 dimension feature vector to overcome the disadvantage of Permutation entropy segmentation error for the similar texture images. They also used Fuzzy C-Means algorithm for the task of cluster texture images. The results show that the method has a good performance on segmenting images with texture distributing uniformly.

Lei Qin et al (2008) presented a novel approach to classify texture collections and the performances of classifying new images using the parameters from the unnoted image correction. The authors also demonstrated that, it is possible to discover texture category from a set of unlabelled images in an unsupervised manner. From their work, it is identified that their approach is robust to significant scale and view point changes in order to achieve the good classification accuracy.

Claudia Chevrefils et al (2009) took the problem of textured image segmentation based on unsupervised scheme with the help of Stochastic Expectation Maximization (SEM). They also applied two dimensional causal nonsymmetrical half plane autoregressive of the textured image as the first stage of the proposed implementation; the SEM procedure is then applied to the set of autoregressive features as the second stage and finally yielding an estimate of the true number of texture classes. They also provided a comparison of the proposal with some previous methods using the same image databases.

Yong Hu and Chun Xia Zhao (2009) proposed a new method for unsupervised texture classification by combining multiscale features and K-Means classifiers. They used heroic features extracted from Gray Level Co-Occurrence Matrix (GLCM) in order to find the optimal window size by taking different window sizes. From their proposed methods, it is found that the achievement of higher classification rate than the single method over fixed window size.

Tassetti et al (2010), textures developed spectral/texture classification schema is compared with the classical approach using only spectral information. An accuracy assessment is carried out which shows that the image data with an accuracy of 80.01% by using texture bands compared to 64% accuracy achieved by using spectral band only. It is understood that the problems in terms of class information extraction especially using pixel based image classification methods in which spatial information existing between a pixel and its neighbors are not used.

Castaello et al (2003) presented a Neuro- fuzzy approach for classification of image pixel into three classes namely contour, regular and texture points. Exploiting the processing capability of a neural network fuzzy classification rules are derived by learning from data and applied to classifying pixels in gray level images. The authors considered the spatial properties of the image features and a multiscale representation of images to derive a proper set of training data. In order to prove the effectiveness of the proposed methodology some experimental results done by them concerned different real world images.

Simon.E et al (2002) compare a number of texture operators that compromise a Gabor filtering stage followed by different types of nonlinear forced Gabor processing. In their work, the authors did not treat the problem of feature selection and feature space dimensionality reduction. Moreover the texture detection capabilities of the operators and their robustness to non texture features are also compared. It is also understood that the operator is the only one that selectively response to texture and does not give false response to non textured features such as object contours.

2.4 APPLICATIONS OF TEXTURE ANALYSIS

Previous section deals with the literature survey on different supervised and unsupervised classification methods. This section renders the different applications by applying supervised and unsupervised classification methods.

Marios Gavrieldes et al (2000) developed a multistage Computer-Aided Diagnosis (CAD) scheme for the automated segmentation of suspicious micro-calcification clusters in digital mammograms. The proposed scheme consisted of three main processing steps. First, the breast region was segmented and its high-frequency content was enhanced using unsharp masking. In the second step, individual micro calcifications were segmented using local histogram analysis on overlapping sub images. The final step clustered the segmented micro calcifications and extracted the following features for each cluster. The results showed a true positive rate of 93.2% and an average of 0.73 false positive clusters per image.

Acha Serrano and Roa (2001) described the algorithm with an aim at separating burned skin from normal skin in burn color images and to classify them according to the depth of the burn. The segmentation procedure consists of an elaborated treatment of color representation, followed by a grayscale segmentation algorithm based on the stack mathematical approach. The proposed algorithm has been developed to be applied to skin wound images, but it works properly as a general segmentation approach. In the classification part, the authors take advantage of color information by clustering, with a vector quantization algorithm.

Qiang et al (2000) presented a generalized statistical texture analysis technique for characterizing and recognizing cervical lesions for colpscopic images. They offered three proposals namely the introduction of generalized texture analysis with the combination of statistical and structural approaches, the introduction of a set of textural measures and the experimental studies with a real image. For classification, they employed the minimum distance classifier with the best discriminant performance of 87.13% by using all 24 features.

Li Wang et al (2001) presented a statistical model supported approach for enhanced segmentation and extraction of suspicious mass areas from mammographic images. In their studies, one type of morphological operation is derived to enhance disease patterns of suspected masses by cleaning up unrelated background clutters, and model-based image segmentation is performed to localize the suspected mass areas using a stochastic relaxation labeling scheme. They also discussed the importance of model selection when a finite generalized Gaussian mixture is employed. The experimental results demonstrate that the proposed method achieves a very satisfactory performance as a preprocessing procedure for mass detection in Computer Aided Design (CAD).

Shuttleworth et al (2002) showed that classification using color texture offers an improvement over classification based solely on grey-level texture using statistical analysis. Worldwide, colorectal cancer is the third most common malignant neoplasm. Automated classification of cytological images of colon tissue samples has been investigated, but diagnosis in all cases still requires human judgment. With the large numbers of cases of colon cancer each year, the workload placed on pathologists is immense. Texture is a powerful discriminating metric and the use of grey-level texture for classification of colon images has been extensively researched. One common technique is the extraction of texture metrics from grey-level co-occurrence matrices. However, using grey-scale images discards information contained in the differences of hue and saturation that may provide further classification information. The authors also presented the findings of an investigation of the discriminating ability of color texture using co-occurrence matrices. Comparisons are made between grey-scale and color texture analysis.

John Filippas et al (2003), proposed the texture analysis techniques which are used to measure certain characteristic of normal and cancer less tissue images with the help of a genetic algorithm for the purpose of maximizing the classification accuracy. In this methodology a total number of 31 images (15 normal and 16 cancerless) are used with the division into two sets containing no common elements. The results are promising with the achievement of 100% accuracy in the classification of the images in the training set and in some cases 91% in that of the test set.

Mark Sheppard and Liwen Shih (2005) presented an efficient integrated images textural analysis and classification of ultrasound images into clusters potentially representing cancerous or normal tissue areas. Their approach is based on Haralick's textural features and the minimum squared error clustering algorithm which offer significant reduction in run time, potentially allowing more accurate and allows the investigation of parameters associated with textural and clustering process.

Poonguzhali and Ravindran (2006) studied the classification of ultrasonic liver images by using texture features extracted from four methods namely Law's method, auto correlation method, Gabor wavelet method and edge frequency method. The features from this method are used to classify three sets of ultrasonic liver images and its suitability in classifying the abnormalities. From their work, it is seen that the Gabor wavelet classifier is better, when compared to other method since it analyzes the method in both time and frequency domain simultaneously.

Alfonso Rojas Dominguez and Asok Nandi et al (2007) presented two new boundary tracing algorithms for segmentation of breast masses is presented. These new algorithms are based on the Dynamic Programming Based Boundary Tracing (DPBT). The DPBT algorithm contains two main steps: (1) construction of a local cost function and (2) application of dynamic programming to the selection of the optimal boundary based on the local cost function. The segmentation results are evaluated with base on the area overlap measure and other segmentation metrics. The DPBT algorithm outperforms the existing algorithm in the way of producing highest segmentation accuracy.

Deepa Lakshmi et al (2007) summarized the various statistical and spectral texture parameters extraction processing, optimal feature technique and automated classification procedures with the use of k-means clustering and neural network based automatic classification. In their work, a multiclass multivariant problem like classification of the liver using ultrasound liver images is automated to eliminate noise and improve the contrast thereby improving the image quality with the rate of 70% using k-means clustering.

Kwitt and Uhl (2008) proposed a set of new wavelet-domain based color-texture features for the classification of zoom-endoscopy images in the field of medical imaging. They extended the concept of classic co-occurrence matrices to capture information between detail-sub band pairs of different color channels. The results show that the proposed features outperform other wavelet-domain based color-texture features in terms of leave one out cross validation accuracy.

William Nailon et al (2008) presented a method to classify areas within Glast Tumor Volume (GTU) and other clinic relevant regions on Computer Tomography(CT) images. Textural features were calculated on regions based extracted within the bladder rectum and a region identification. For that the sequential forward method was used to reduce the feature set. The results demonstrate, the significant sensitivity of the reduced feature set for classification of any orthogonal CT images.

Mascaro et al (2009) presented the use of three texture descriptors for breast tissue segmentation purposes: the Sum Histogram, the Gray Level Co-Occurrence Matrix (GLCM) and the Local Binary Pattern (LBP).Tissue classification in mammography can help the diagnosis of breast cancer by separating healthy tissue from lesions. A modification of the LBP is also proposed for a better distinction of the tissues. In order to segment the image into its tissues, these descriptors are compared using a fidelity index and two clustering algorithms: K-Means and SOM (Self-Organizing Maps).

Jumali et al (2009) reviewed the benefit of network-based in term of interaction data for classification in identification of class cancer. The advent of high throughput techniques such as micro array data enabled researchers to elucidate process in a cell that fruitfully useful for pathological and medical. For such opportunities, micro array gene expression data have been explored and applied for various types of studies. Unfortunately, since gene expression data naturally have a few of samples and thousands of genes, this leads to a biological and technical problems. Thus, the availability of artificial intelligence techniques couples with statistical methods can give promising results for addressing the problems. These approaches derive two well known methods: supervised and unsupervised. From their studies, it is identical that whenever possible, these two superior methods can work well in classification and clustering in term of class discovery and class prediction.

Yanfen Guo et al (2009) studied and analyzed the image features by using the combined supervised classification method with knowledge classification method in order to extract the forest investigation. Applying the fused image with the multispectral band, the shadows are separated with the threshold of the texture images, where the pixels are recovered with the method of histogram matching. The overall accuracy reaches 84.6% and smallest forest accuracy is 78.99%.

Claudia Chervefils et al (2009) presented a unified frame work for automatic segmentation of intervetal disk of scoliotic spine from magnetic resonance images (MRI). The methods exploit a combination of statistical and spectral texture features to discriminate closed regions. Statistical texture features based on the GLCM are also extracted from each closed regions. Each closed region is represented by three feature configuration namely all statistical and spectral texture features transformation, Haralick statistical and combined statistical and Haralick transformation. The results suggest that the selected texture feature classification can be contributed the problem of over segmented by successfully discriminative intervetal disk of scoliotic spine from Magnetic Resonance (MR) images.

Deepa Sankar and TessammanThomas (2009) presented a fast fractal method to model breast back ground regions based on entropy for the detection of breast cancer. In their work, the domain pool for searching the matching domain is chosen based on entropy. This reduced the encoding time by a factor of 3.12 when compared with the conventional fractal encoding method which searched the entire domain pool for searching domain with the reduction rate of 85% for the 28 abnormal images used.

Zhimin Lan et al (1995) developed a technique that analyzes patterns and quantifies the degree of speculation present. The approach involves two process namely automatic lesion extraction using region growing and feature extraction using radial edge gradient analysis. The performance of each of the two measures of speculation was tested on a database of 95 mammography masses using ROC analysis that evaluates their individual ability to determine the likelihood of malignancy of masses. The maximum value of one of the two speculation measures (FWHM) from the four neighborhoods yielded a very good classification accuracy of mammography mass lesions.

Sahiner et al (1996) introduced a new Rubber Band Straightening Transform (RBST) for characterization of mammography masses as malignant or benign. The RBST transforms a band of pixels surrounding a segmented mass onto the Cartesian plane (the RBST image). The border of a mammography mass appears approximately as a horizontal line, and possible speculations resemble vertical lines in the RBST image. Texture features extracted from spatial gray-level dependence matrices and run-length statistics matrices were evaluated for different regions and representations. The ROC was used to evaluate the classification accuracy, which is seemed to be more effective than those extracted from the original images.

Mudigonda et al (2000) studied the computer aided classification of gradient-based and texture-based features. Features computed based on gray-level co-occurrence matrices (GLCMs) are used to evaluate the effectiveness of textural information possessed by mass regions in comparison with the textural information present in mass margins. A total of 54 images (28 benign and 26 malignant) containing 39 images from the Mammographic Image Analysis Society (MIAS) database and 15 images from a local database have been taken for the test. The best benign versus malignant classification of 82.1%, with an area (A_z) of 0.85 under the Receiver Operating Characteristics (ROC) curve was obtained with the images from the MIAS database by using GLCM-based texture features computed from mass margins. The results show that the proposed method has very good classification accuracy.

Verma and Zakos (2001) presented a system based on fuzzy-neural and feature extraction techniques for detecting and diagnosing micro calcifications patterns in digital mammograms. They investigated and analyzed a number of feature extraction techniques and found that a combination of three features (such as entropy, standard deviation and number of pixels) is the best combination to distinguish a benign micro calcification pattern from one that is malignant. This method is an easy-to-use intelligent system that gives the better classification accuracy in digital mammograms.

Hadjiiski et al (2001) developed a new classification scheme to classify mammographic masses as malignant and benign by using interval change information. The masses on both the current and the prior mammograms were automatically segmented using an active contour method. From each mass, 20 Run Length Statistics (RLS) texture features, 3 speculation features, and 12 morphological features were extracted. Stepwise feature selection and linear discriminant analysis classification were used to select and merge the most useful features. The information on the prior image significantly (p=0.015) improved the accuracy for classification of the masses.

Sheila Timpand and Nico Karssemeijer (2001) developed computer aided diagnosis (CAD) techniques to study interval changes between two consecutive mammographic screening rounds. The goal of their work was to improve the detection method and to improve the temporal information in the CAD program apart from developing a regional registration technique, which links a suspicious location on the current mammogram with a corresponding location on the prior mammogram. Temporal features are obtained by combining the feature values from both regions. The analysis shows an improvement in detection performance with and without the use of temporal features.

Zhimin Huo et al (2000) proposed to identify computer-extracted, mammography parenchyma patterns that are associated with breast cancer risk. They extracted 14 features from the central breast region on digitized mammograms to characterize the mammography parenchyma patterns of women at different risk levels. The features were used to distinguish mammography patterns seen in low-risk women from those, who inherited a mutated form of the gene, which confers a very high risk of developing breast cancer. The results show that women at high risk tend to have dense breasts and their mammography patterns tend to be coarse and low in contrast.

Brijesh Verma and Zhang (2007) proposed a neural-genetic algorithm for feature selection to classify micro calcification patterns in digital mammograms, which aims at developing a step-wise algorithm to find the best feature set and a suitable neural architecture for micro-calcification classification. The obtained results show that the proposed algorithm is able to find an appropriate feature subset, which also produces a high classification rate.

Mavroforakis et al (2002) presented an analytical approach to texture classification, combined with qualitative descriptive diagnostic data. For qualitative data, a statistical approach was applied in detailed clinical findings and texture-related features were established as of most importance during the diagnostic assertion process. A complete set of textural feature functions in multiple configurations and implementations was applied to a large set of digitized mammograms, in order to establish the discriminating value and statistical correlation with qualitative texture descriptions of breast mass tissue. The results shows that the optimal classification accuracy rates reached 81.5% for texture-only classification and 85.4% with the introduction of patient's age as an example of hybrid approaches.

Keir Bovis et al (2000) studied the identification of masses in digital mammograms using texture analysis. A number of texture measures are calculated for bilateral difference images showing regions of interest. The measurements are made on co-occurrence matrices in different directions. The results are generated on the average recognition rate over these folds on correctly recognizing masses and normal regions. It is concluded, that the recognition rates of 77% correct recognition and an area under the ROC curve value.

Bhangale Tushar et al (2000) employed a Gabor filter bank for texture analysis of mammograms to detect micro-calcifications. A subset of the Gabor filter bank with a certain central frequency and different orientations is used to obtain the Gabor-filtered images, which are then subjected to a histogram based threshold to obtain binary images. From their studies it is identified that the ratio of the micro-calcification clusters correctly detected to the total number of clusters present in the image and false-positive rate is the number of detections in which normal breast parenchyma is classified as a micro calcification cluster.

Grey level co-occurrence features are one of the most powerful feature sets available for texture analysis. However, the moving window commonly employed to define the statistical scale at which the co-occurrence matrix is obtained assumes spatial stationary of the underlying random field. This assumption is inappropriate in the case of natural images and may result in the mixing of different structures at various positions that can yield misleading features, affecting any subsequent analysis or classification. To minimize this problem, Fernando Bello and Richard (1996) presented a method for obtaining co-occurrence features from the irregular tessellation of an image. Such tessellation is considered to be the result of a filtering or pie-segmentation step guaranteeing a certain degree of homogeneity within each tessellation element, and thus offering a more optimal statistical scale at each location in the image. Experimental results and a comparison between features obtained from various irregular and square tessellation elements in a set of natural texture images are presented. The authors also show that the features obtained have a similar behavior to those generated from a traditional square window.

Xinbo Gao et al (2002) presented a novel clustering algorithm supervised by statistical tests. It deals with three key problems in data analysis, cluster tendency, cluster analysis and cluster validity, simultaneously. So, it provides an effective tool to analyze the validity of pattern unsupervised classification, especially in the case of large number of samples. The experimental result illustrates its effectiveness in classifications.

Jaime Melendez and Domenec Puig (2010) proposed a new efficient methodology for unsupervised image segmentation based on texture. It takes advantage of a supervised pixel-based texture classifier trained with feature vectors associated with a set of texture patterns initially extracted through a clustering algorithm. Therefore, the final segmentation is achieved by classifying each image pixel into one of the patterns obtained after the previous clustering process. Multi-sized evaluation windows following a top-down approach are applied during pixel classification in order to improve accuracy. The proposed technique has been experimentally validated on MeasTex, VisTex and Brodatz compositions, as well as on complex ground and aerial outdoor images. Comparisons with state-of the-art unsupervised texture segmenters are also provided.

The above literature survey was presented with an extensive review on various existing classification methods addressing different texture parameters and their applications. Still there is a need for further research in finding improved algorithms to increase the classification accuracy in texture analysis and its application. Based on the above literature survey, this monograph focuses on the reliability, classification accuracy, Transmission Time, kurtosis and skewness in supervised and unsupervised classification methods.

The supervised and unsupervised classification of images using textural features is proposed in Chapter 3.

CHAPTER 3
TEXTURE ANALYSIS METHODS FOR FEATURE EXTRACTION

3.1 INTRODUCTION

In the previous Chapters, introduction and the literature survey in the area of texture analysis methods, related classification techniques, applications and work done so far to improve the classification accuracy have been discussed. Texture is an important spatial feature, useful for identifying objects or regions of interest in an image. Texture is the term used to describe the surface of a given phenomenon in an image. Texture analysis plays an increasingly important role in digital image processing and pattern recognition and is widely applied to the processing and interpretation of remotely sensed data, biomedical and microscopic cell images, where texture information is sometimes the only way to characterize a digital image. Texture Spectrum Operator (TSO) method is statistical approach to texture analysis where the local texture for a given pixel and its neighborhood is characterized by the corresponding texture unit and an image can be characterized by its texture spectrum which is the occurrence frequency function of all texture units within the image. The local binary pattern operator (LBP) is defined as a gray scale invariant texture measure derived from a general definition of texture in local neighborhood. In order to overcome the problems existing in Texture Spectrum Operator (TSO), and Local Binary Pattern Operator (LBP), there is a high requirement to go for a better classification method. Entropy Based Local Descriptor (EBLD) method may be highly suitable for the texture measure to quantify the smoothness of an image texture, since Entropy does not depend on actual values in texture. This chapter describes the different supervised classification methods that include Texture Spectrum Operator method (TSO) and Local Binary Pattern Operator (LBP). With the proposed Entropy Based Local Descriptor (EBLD) method to extract the texture features. In addition, applying the proposed Partition Clustering Method (PC) and Fuzzy Partition Clustering Method (FPC) in the unsupervised classification methods to increase the classification accuracy is also described.

The objective of this chapter is to study the concept of supervised and unsupervised classification using texture features to demonstrate the potential usefulness of the proposed methods in texture analysis.

3.2 NEED FOR SUPERVISED AND UNSUPERVISED CLASSIFICATION

Methods of texture analysis are usually divided into two major categories (Dong Chen He and Li Wang 1991). The first is the structural approach, where texture is considered as a repetition of some primitives, with a certain rule of placement. The traditional Fourier spectrum analysis is often used to determine the primitives and placement rule. Several authors have applied this method to texture classification and texture characterization with a certain degree of success (Haralick 1979). Problems may be encountered in practice in identifying the primitives and the placement rule in natural images. The second major approach in texture analysis is the statistical method with the aim at characterizing the stochastic properties of the spatial distribution of gray levels in an image. So some statistical methods of texture analysis are needed to focus on texture characterization and discrimination with or without the prior knowledge about the texture features. So it is inevitable to study the different supervised and unsupervised classification methods.

3.3 EXISTING SUPERVISED CLASSIFICATION METHODS

Previously for achieving better classification on texture analysis, the two important supervised classification methods used are Texture Spectrum Operator (TSO) and Local Binary Pattern Operator (LBP).

3.3.1 Local Binary Pattern Operator Method (LBP)

The original LBP method is a complementary measure for local image contrast. The LBP method code is to be produced by multiplying the threshold values by weights given by powers of two. Normally LBP method is invariant to any monotonic transformation of the gray scale and it's quick to compute with larger neighborhoods, the number of possible local binary pattern method codes increased exponentially. This can be avoided to some extent by considering only a subset of that codes one approach is to use so called uniform patterns representing the statistically most LBP method code. With them the size of the feature histogram generated by a LBP method can be reduced without significant loss in its discrimination.

The Figure 3.1 shows the computation of LBP that includes the pixel, threshold and weights that are obtained by operating with eight neighboring pixels using the center as a threshold.

7	1	12
2	5	5
5	3	0

(a) Pixels(example)

1	0	1
0		1
1	0	0

(b)Threshold

1	0	4
0		16
32	0	0

c) Weights

Figure 3.1 Computation of Local Binary Pattern (a) Pixel (b) Threshold and (c) Weights

LBP method is described with 2^8 possible texture units. The texture unit is obtained by applying the threshold operation using the following rule

$$\mathrm{Ei} = \begin{cases} 0, V_i < V0 \\ 1, V_i >= V0 \end{cases} \qquad (3.1)$$

where, V_0 is the center pixel. The LBP operator method is determined as

$$\mathrm{LBP} = \sum_{I=1}^{8} E_i \times 2^{i-1} \qquad (3.2)$$

LBP operator method does not take into account the contrast of texture which is the measure of local variations present in an image and is important in the description of some textures. Texture Spectrum Operator (TSO) is similar to LBP operator method but it uses three levels that is, two thresholds instead of two levels used in LBP operator method. This leads to a more efficient representation and implementation than with LBP operator method and three level operators does not perform better than LBP method.

3.3.2 Texture Spectrum Operator (TSO)

The local texture for a given pixel and its neighborhood is characterized by the corresponding texture unit and an image can be characterized by its texture spectrum in statistical approach for texture analysis which is the occurrence frequency function of all texture units within the image. In a square raster digital image each pixel is surrounded by eight neighborhood pixels. The local texture information for a pixel can be extracted from a neighborhood of 3×3 pixels which is denoted by a set containing nine elements V= $\{v_0, v_1 \ldots\ldots v_8\}$, Where v_0 represents the intensity value of the central pixel and vi{i=1,2,….8} is the intensity value of the neighboring pixel 'i' to define the corresponding texture unit by a set containing eight elements. Texture Unit (TU) = {E1,E2,…E8}where E_I{i=1,2,…8} is determined by the formula

$$Ei = \begin{cases} 0 \text{ if } Vi < Vo \\ 1 \text{ if } Vi = Vo \\ 2 \text{ if } Vi > Vo \end{cases} \qquad (3.3)$$

For i =1, 2…8 and the element Ei occupies the same position as the pixel i. As each element of texture unit has one of three possible values with the combination of all eight elements results in 3^8=6561 possible texture units in total. Since there are three comparison levels (<, =,>) and have called this method as Texture Spectrum Operator. For N= 3, the combinations of all the elements results in 3^8=6561 possible texture units.

3.4 EXISTING UNSUPERVISED CLASSIFICATION METHODS

The existing unsupervised classification method to extract the texture features are namely K-Means clustering and Fuzzy C-Means clustering method.

3.4.1 K-Means Clustering

The K-means clustering is used in computer vision as a form of image segmentation. The results of the segmentation are used to aid border detection and object recognition. In this context, the standard Euclidean distance is usually insufficient in forming the clusters. Instead, a weighted distance measure utilizing pixel coordinates, RGB pixel color and/or intensity, and image texture is commonly used.

The most common algorithm uses an iterative refinement technique. Due to its ubiquity it is often called the *k*-means algorithm. Given an initial set of k means $\mathbf{m}_1^{(1)},\ldots,\mathbf{m}_k^{(1)}$, which may be specified randomly or by some heuristic, the algorithm proceeds by alternating between two steps.

- **Assignment step**: Assign each observation to the cluster with the closest mean (i.e. partition the observations according to the Voronoi diagram generated by the means).
- **Update step**: Calculate the new means to be the centroid of the observations in the cluster.

Given a set of observations ($\mathbf{x}_1$, $\mathbf{x}_2$, …, $\mathbf{x}_n$), where each observation is a d-dimensional real vector, k-means clustering aims to partition the n observations into k sets ($k \leq n$) $\mathbf{S} = \{S_1, S_2, \ldots, S_k\}$ so as to minimize the within-cluster sum of squares (WCSS), where, $\boldsymbol{\mu}_i$ is the mean of points in S_i.

The algorithm is deemed to have converged when the assignments no longer change. The demonstration of the standard algorithm is as follows.

1) k initial "means" (in this case k=3) are randomly selected from the data set (shown in color).
2) k clusters are created by associating every observation with the nearest mean.
3) The centroid of each of the k clusters becomes the new means.
4) Steps 2 and 3 are repeated until convergence has been reached.

As it is a heuristic algorithm, there is no guarantee that it will converge to the global optimum, and the result may depend on the initial clusters. As the algorithm is usually very fast, it is common to run it multiple times with different starting conditions. However, in the worst case, *k*-means can be very slow to converge: in particular it has been shown that there exist certain point sets, even in 2 dimensions, on which *k*-means takes exponential time, that is $2^{\Omega(n)}$, to converge.

These point sets do not seem to arise in practice: this is corroborated by the fact that the smoothed running time of *k*-means is polynomial.

Drawbacks of K-means clustering

The two key features of *k*-means which make it efficient are often regarded as its biggest drawbacks:

- Euclidean distance is used as a metric and variance is used as a measure of cluster scatter.
- The number of clusters *k* is an input parameter: an inappropriate choice of *k* may yield poor results. That is why, when performing k-means, it is important to run diagnostic checks for determining the number of clusters in the data set.

A key limitation of *k*-means is its cluster model. The concept is based on spherical clusters that are separable in a way so that the mean value converges towards the cluster center. The clusters are expected to be of similar size, so that the assignment to the nearest cluster center is the correct assignment. When for example applying *k*-means with a value of $k = 3$ onto the well-known Iris flower data set, the result often fails to separate the three Iris species contained in the data set. With $k = 2$, the two visible clusters will be discovered, whereas with $k = 3$ one of the two clusters will be split into two even parts. In fact, $k = 2$ is more appropriate for this data set, despite the data set containing 3 classes. As with any other clustering algorithm, the *k*-means result relies on the data set to satisfy the assumptions made by the clustering algorithms. It works very well on some data sets, while failing miserably on others.

3.4.2 Fuzzy C-Means Clustering (FCM)

The FCM algorithm is one of the most widely used fuzzy clustering algorithms. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. The FCM algorithm attempts to partition a finite collection of elements $X= \{x_1, x_2, \ldots, x_n\}$ into a collection of c fuzzy clusters with respect to some given criterion. Given a finite set of data, the algorithm returns a list of c cluster centers V, such that

$V = v_i$, i =1, 2, ... , and a partition matrix U such that

$U = u_{ij}$, i =1, ..., c, j =1... n

where u_{ij} is a numerical value in [0, 1] that tells the degree to which the element x_j belongs to the i-th cluster.

The following is a FCM algorithm procedure in steps..

Step 1: Select the number of clusters c ($2 \leq c \leq n$), exponential weight μ($1 < \mu < \infty$), initial partition matrix U^0, and the termination criterion ϵ. Also, set the iteration index l to 0.

Step 2: Calculate the fuzzy cluster centers $\{v_i^1 \mid i=1, 2, \ldots, c\}$ by using U^1.

Step 3: Calculate the new partition matrix U^{l+1} by using $\{v_i^1 \mid i=1, 2, \ldots, c\}$.

Step 4: Calculate the new partition matrix $\Delta = \|U^{l+1} - U^1\| = \max_{i,j} |u_{ij}^{l+1} - u_{ij}^1|$.

If $\Delta > \sum$, then set $l = l + 1$ and go to step 2. If $\Delta \leq \epsilon$, then stop. This method is frequently used in pattern recognition.

3.5 PROPOSED SUPERVISED AND UNSUPERVISED CLASSIFICATION METHODS

To achieve high classification accuracy using supervised and unsupervised methods the Entropy Based Local Descriptor (EBLD) with Partition Clustering Method (PCM) and Fuzzy Partition Clustering Method (FPCM) are studied in this section.

3.5.1 Proposed Entropy Based Local Descriptor (EBLD)

Entropy based local descriptor is a measure of information content. It measures the randomness of intensity distribution and is largely depending on the probability. It is represented by

$$\text{EBLD}=\sum_{I=1}^{8} P(i)\log p(i) \qquad (3.4)$$

Such a matrix corresponds to an image in which there are no preferred gray level pairs for the distance vector d. Entropy Based Local Descriptor is highest when all entries in P (i) are of similar magnitude, and small when the entries in P[i,j] are unequal. Using P (i) the entropy value has been found out at every smooth region. Moreover the P(i) determines the EBLD value. Entropy Based Local Descriptor operator calculates the brightness in a local region of the picture. The EBLD value is maximum when the brightness in a local region of the picture. EBLD is a measure of histogram uniformity.

Entropy Based Local Descriptor is a texture measure widely used to quantify the smoothness of image texture. Entropy Based Local Descriptor does not depend on actual values in texture but it does depend on the smoothness of texture. Entropy Based Local Descriptor measures the randomness of the elements of a matrix, when all the elements of the matrix are maximally random Entropy Based Local Descriptor has its highest value, so a homogenous image has lower EBLD than an in homogeneous image. The EBLD defined as the average number of binary symbols necessary to code a given input given the probability of that input appearing an a stream. High EBLD is associated with a high variance in the pixel values, while low EBLD indicates that the pixel values are fairly uniform. When applied to groups of pixels within the source images, EBLD provides a way to compare regions from the different source images. EBLD operator described with 2^8 possible textures and the proposed supervised algorithm is given as follows.

3.5.1.1 Proposed EBLD Supervised Texture Classification Algorithm

Texture is the most important visual cue in identifying these types of homogeneous regions. The goal of texture classification is to produce a classification map of the input image where each uniform textured region is identified with the texture class it belongs to. The entire process for performing the different classification methods for the proposed supervised classification algorithm is presented below.

Step 1: Each image consists of 256 ×256 pixels with 64 normalized gray levels.

Step 2: A sample sub image of 30 ×30 pixels was selected within each texture. Using a window of 30 ×30 pixels together with a step of two pixels in the row and column.

Step 3: The texture spectrum was calculated within a window of 30×30 pixels. Calculate the smooth regions using P(i).

Step 4: The minimum distance decision rule was used and the integrated absolute difference between two texture spectra was considered as the distance between them

$$D(i) = \sum_{j=0}^{\text{Range}} \left|(w(j) - S(i,j)\right| \qquad (3.5)$$

where D (i) denotes the absolute difference between the texture spectrum of target image and the texture spectrum of a sample sub image I, W (j) represents the occurrence value of texture unit j in the texture spectrum of the window considered with the target image, S (i,j) represents the occurrence value of texture unit j in the texture spectrum of the sample sub image i.

Step 5:. The percentage of classification are obtained by number of pixels misclassified divided by the total number of pixels in the image. The steps for the supervised classification are shown in Figure 3.2

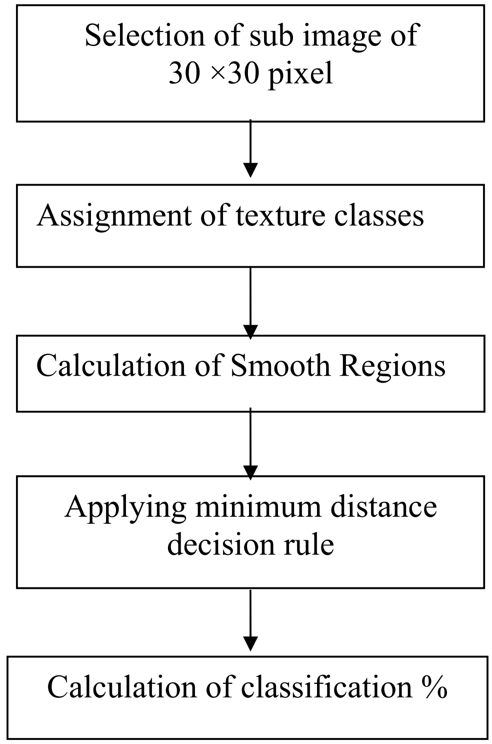


Figure 3.2 Proposed EBLD Supervised Classification Method

3.5.2 Proposed Partition Clustering (PC) Method

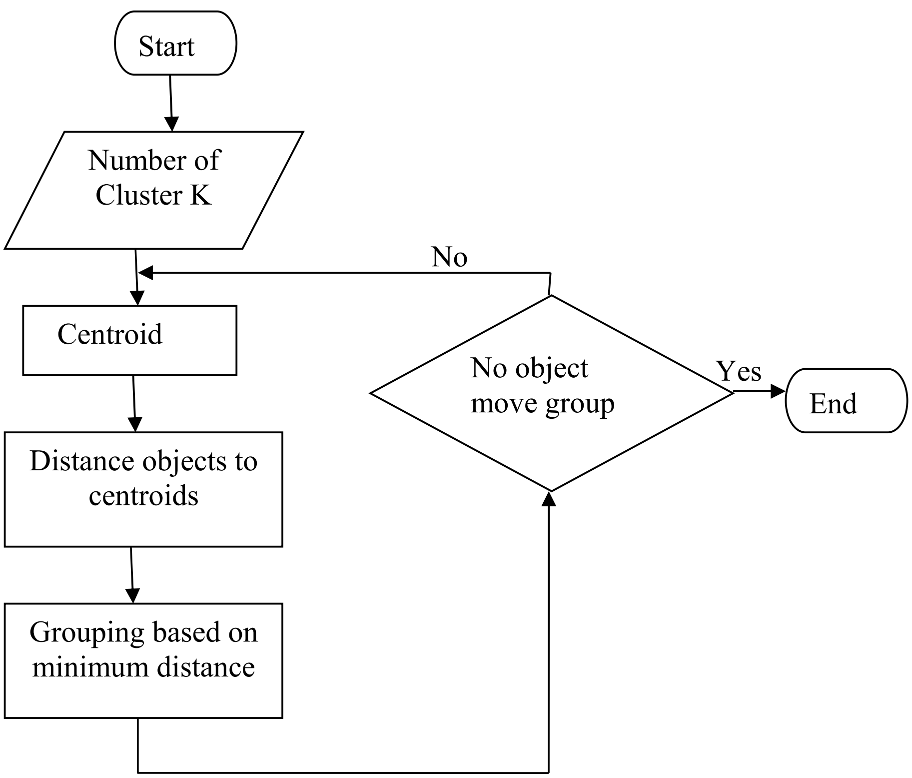


Figure 3.3 Flow Chart of Partition Clustering Method

The aim of the Partition Clustering (PC) Method is to classify or to group the objects based on attributes/features into K number of group where K is positive integer number. The grouping is done by minimizing the sum of squares of distances between data and the corresponding cluster centroid. The Partition Clustering method is shown in Figure 3.3.

The proposed Partition Clustering (PC) Method is explained as follows.

Step 1: Begin with a decision on the value of K= number of clusters

Step 2: Put any initial partition that classifies that data in to K clusters. Then the training samples may be assigned randomly or systematically as per the following.

- Take the first K training samples as single element cluster.
- Assign each of the remaining (N-k) training samples to the cluster with the nearest centroid. After each assignment, recompute the centroid of the gaining cluster.

Step 3: Take each sample in sequence and compute its distance from the centroid of each of the clusters. If a sample is not currently in the cluster with the closest centroid switch this sample to that cluster and update the centroid of the cluster gaining the new sample and the cluster losing the sample and repeat step 3 until convergence is achieved.

If the number of data is less than the number of clusters then each data is to be assigned as the centroid of the cluster. Each centriod will have a cluster number. If the number of data is bigger than the number of cluster, for each data, the distance calculation to all centroid and get the minimum distance.

3.5.3 Proposed Fuzzy Partition Clustering (FPC) Method

Clustering involves the task of dividing data points into homogeneous classes or clusters so that items in the same class are as similar as possible and items in different classes are as dissimilar as possible. Clustering can also be thought of as a form of data compression, where a large number of samples are converted into a small number of representative prototypes or clusters. Depending on the data and the application, different types of similarity measures may be used to identify classes, where the similarity measure controls how the clusters are formed.

In non-fuzzy, data is divided into crisp clusters, where each data point belongs to exactly one cluster. In the proposed Fuzzy Clustering (FC), the data points can belong to more than one cluster, and associated with each of the points are membership grades which indicate the degree to which the data points belong to the different clusters.

The entire process for performing the different classification methods for the proposed unsupervised classification algorithm is presented below.

Input : Any texture image to be classified with size M×N where $M, N \geq 30$

Output : The regions classified in the image.

Step 1 : Take a window of size 30 x 30 in the image region from the input image.

Step 2 : Take a sub-window of size 3 x 3 in the window image.

Step 3 : The central pixel of the matrix will be assigned by the corresponding entropy based local descriptor number. Compare all pixel values of the window with the center pixel.

Step 4 : If pixel value less than center pixel value, the difference is marked 0. If pixel value greater than or equal to center pixel value, note the difference is 1.

Step 5 : Form a eight digit binary value from the noted difference values (totally 8 pixels except center pixel). Find the gray value of the eight digit binary value. Replace the pixel of the window image using the gray value.

Step 6 : Identifying the different regions and move the sub-window towards the whole window image using overlapping method and find the gray value of the all sub-windows using above mentioned steps and replace the all pixels of the window image using these gray values.

Step 7 : Label the different clusters and Color the different labels.

Step 8 : Calculate the classification accuracy based on the number of classes included for the test and how many classes get assigned. The classification accuracy is the ratio between the numbers of samples misclassified to the number of samples used.

The Figure 3.4 illustrates the FPC unsupervised classification method. Assignment of texture classes has been done by means of comparing the pixel values of the window with the center pixel. The selection of centre pixel is needed here because of knowing the difference of pixel value to replace the pixel of the window image using the gray value.

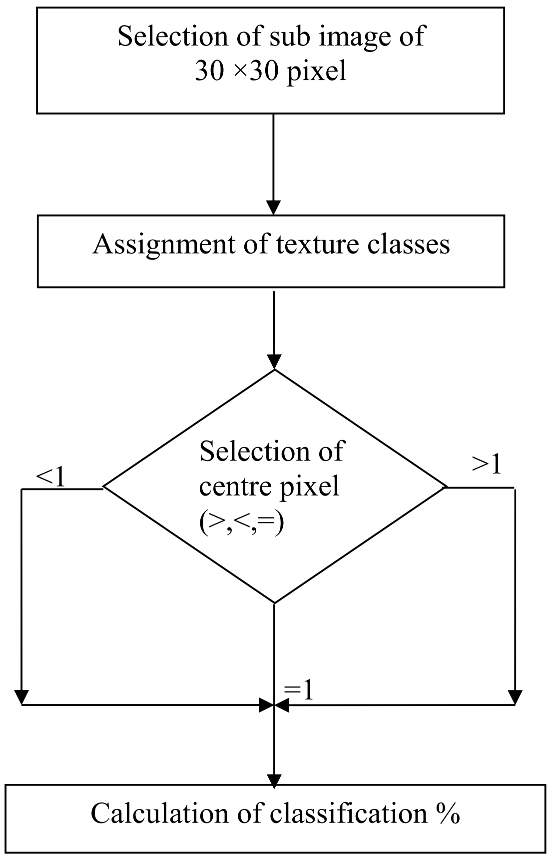


Figure 3.4 Proposed Fuzzy Partition Clustering (FPC) Unsupervised Classification Method

3.6 RESULTS AND EVALUATION

3.6.1 Simulation Study

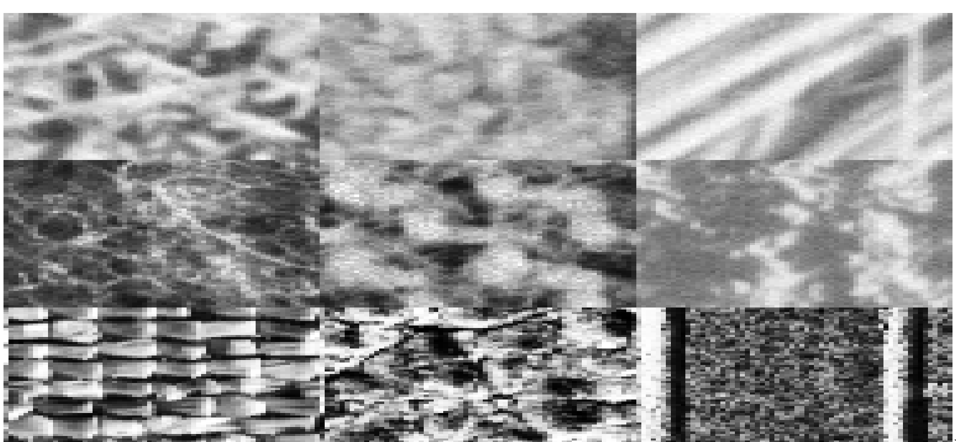
This work is simulated using a math lab Simulator. To evaluate the performance of the proposed method in a realistic scenario, the images from Brodatz database are taken randomly. The defined set of texture unit describes the local texture aspects for a given pixel. That is the relative gray level relationship between the central pixel and its eight neighbors. Thus the statistics of the frequency of occurrence of all the texture units over a large region of an image will reveal the texture information globally.

A sample textured image collected from Brodatz album and its corresponding spectrum of TSO, LBP and EBLD have been taken for the evaluation. Texture classification problem is to assign each possible region or a pixel in the textured image to a known proper texture class. The classification is based on the simple distance function. The computed global descriptors namely with TSO, LBP and EBLD of the known texture classes have been used as the training set of feature vectors. The feature vectors are used during classification of target images having unknown distribution of the known textures. The target image has been classified using TSO, LBP and EBLD.The classification of texture image in the input mosaic image is calculated by the following equation.

Classification Percentage= (number of pixels classified in a texture image / total number of pixels in a texture image)*100

Texture has been one of the most important characteristics which have been used to classify and recognize objects and scenes. It can be characterized by textural primitives as unit elements and neighborhoods in which the organization and relationships between the properties of these primitives are defined. The classification for unsupervised Texture Segmentation and Classification based on features extracted from EBLD have also been done using Partition Clustering (PC) and Fuzzy Partition Clustering (FPC). The texture mosaic images are reconstructed using Entropy operator. The EBLD values are extracted from the entropy image. This method is tested with different texture databases. The proposed EBLD approach gives good accuracy, when compared with other texture feature methods. The percentage of classification for the entropy based operator is 70.88%. But the classification of segmented image has a higher value of 94.12%, when the proposed EBLD is used. In order to evaluate the performance of the LBP, Supervised TSO and EBLD in texture characterization and classification, several experimental studies have been carried out on nine of Brodatz's natural images. These images were selected because they are broadly similar to one another and also that they resemble parts of remotely sensed images.

A target image is created of size 256 ×256 with nine component texture regions each of size 128 ×128 collected from Brodatz textural album. Gray level in the target image ranges from 0-255.The Figure 3.5 shows the Target image for supervised texture classification, where nine classes are represented by nine different colors and classification percentage of each image. Thus the proposed EBLD are applied to texture classification problem and the percentage of correct classification is claimed to be more than 94% for supervised texture classification.



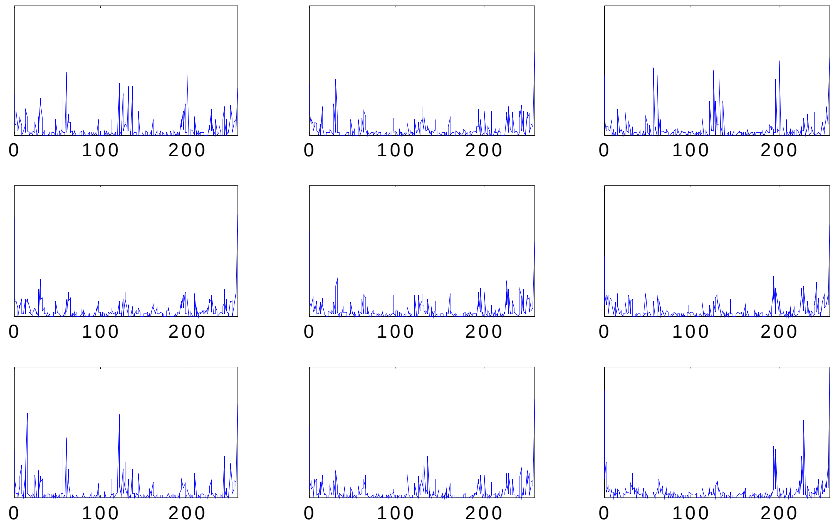


Figure 3.6 Texture Spectrum of Nine images using LBP
(X axis - Texture Unit Number Y axis – Frequency of occurrences)

Figure 3.7 Supervised Classifications of the Nine Input Classes using LBP

Figures 3.6 and 3.7 shows the Texture Spectrum of nine images using LBP operator and the representation of output of nine classes in gray levels respectively. Figures 3.8 and 3.9 show Texture Spectrum of nine images using TSO method and the representation of the nine classes in gray levels respectively. Figure 3.10 and 3.11show Texture Spectrum of nine images using EBLD and the representation of the nine classes in gray levels respectively.

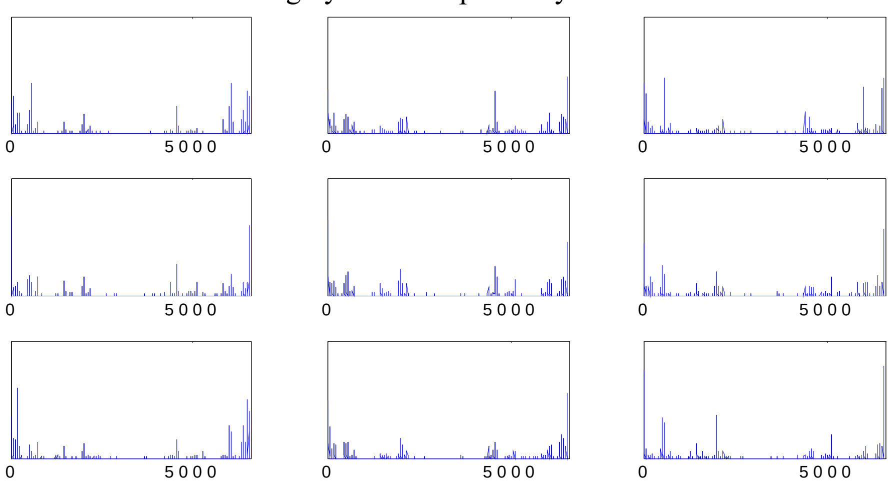


Figure 3.8 Texture Spectrum of Nine Images using Texture Spectrum Operator (TSO)
(X axis - Texture Unit Number Y axis – Frequency of occurrences)

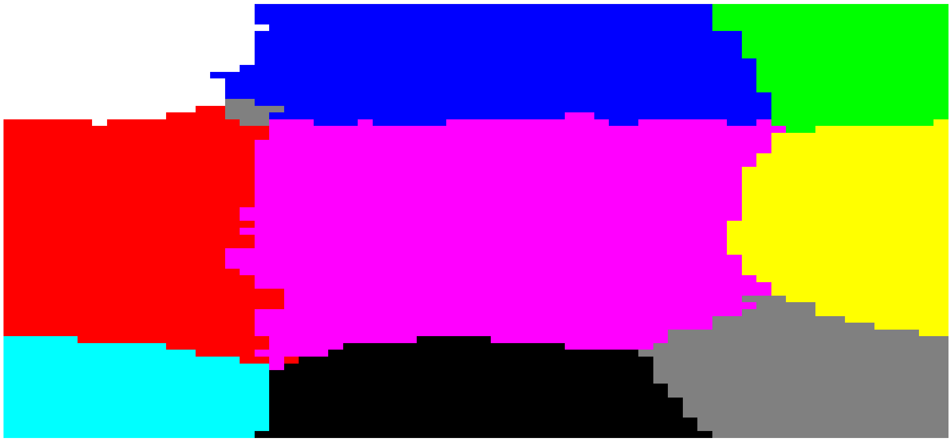


Figure 3.9 Supervised classifications of the nine input classes using TSO

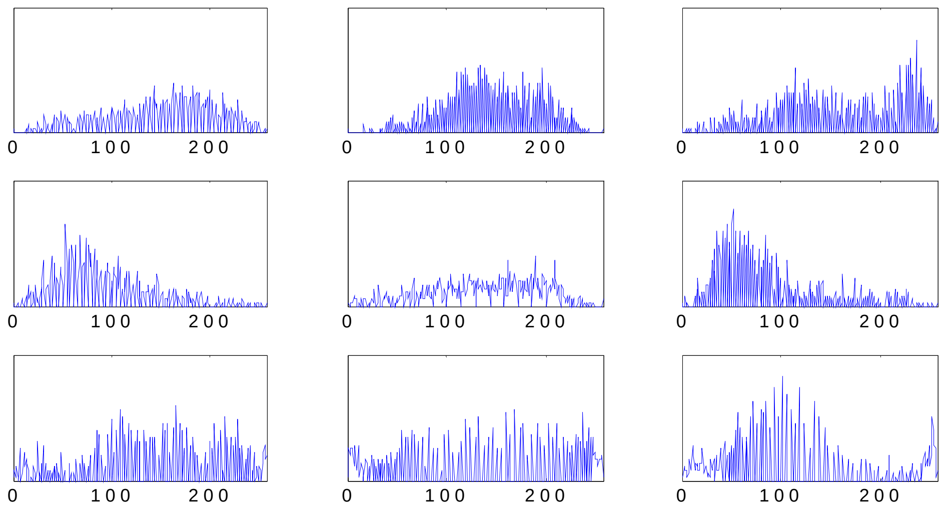


Figure 3.10 Texture Spectrum of Nine Images using EBLD
(X axis - Texture Unit Number Y axis – Frequency of occurrences)

Figure 3.11 Supervised classification of the nine input classes using EBLD

The comparative analysis has been made and it is concluded that the EBLD renders better classification accuracy and is shown in Figure 3.12.

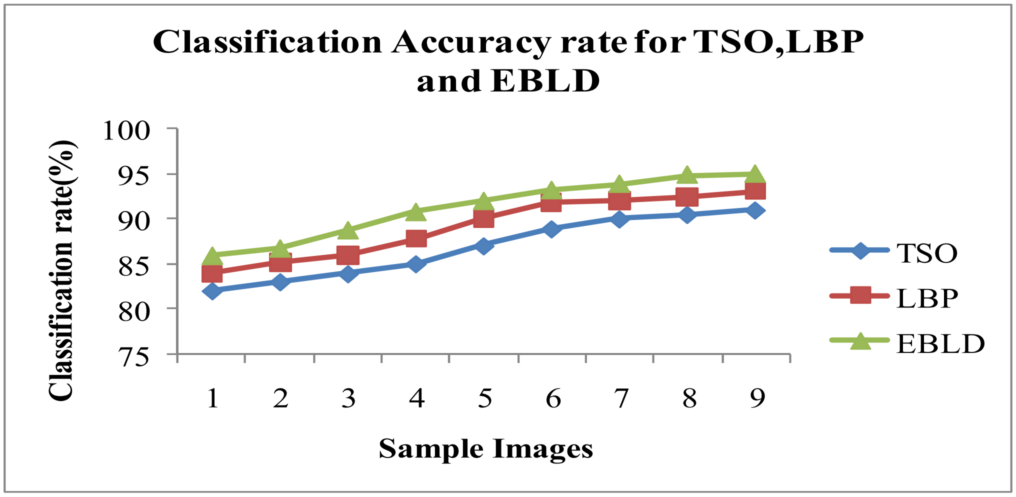


Figure 3.12 Classification Accuracy Rate for TSO, LBP and EBLD

Experimenting with the nine images and applying in the different classification techniques results in supervised classification are noted in the Table 3.1.

Table 3.1 Classification Results in Supervised Classification

Images	Texture Spectrum Operator(TSO)	Local Binary Pattern Operator(LBP)	Entropy Based Local Descriptor method(EBLD)*
D1	82.00	84.00	86.00
D 2	83.00	85.2	86.8
D 3	83.89	86.00	88.8
D 4	85.00	87.8	90.8
D 5	87.00	90.02	92.00
D 6	88.9	91.8	93.2
D 7	90.00	92.00	93.89
D 8	90.5	92.35	94.89
D 9	91.00	93.00	95.00

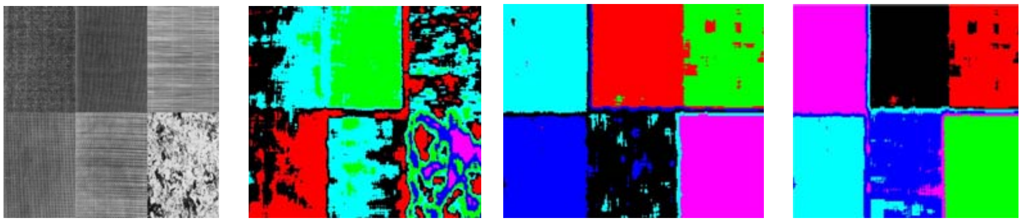
* Proposed

3.6.2 Discussion

This section presents the performance results of the proposed methods compared with various existing methods obtained through simulation. The results are measured in terms of classification accuracy and time consumption.

A target image is created of size 256 ×256 with nine component texture regions each of size 128 ×128 collected from Brodatz textural album. Gray level in the target image ranges from 0-255.Thus the proposed supervised classification methods like entropy based local descriptor and LBP are applied to texture classification problem and the percentage of correct classification for EBLD method is claimed to be more than 94% for supervised texture classification.

In order to evaluate the Partition Clustering (PC) Method classification method with the TSO, LBP and EBLD, the same set of images have been taken. The input images respectively D17, D21, D49, D53, D55 and D60 are shown in Figure 3.13(a). The output image after applying unsupervised FCM classification methods in TSO, LBP and EBLD are shown in Figure 3.13(b) to (d) and the output images after applying unsupervised Partition Clustering classification methods in TSO, LBP and EBLD are shown in Figure 3.14 (b) to (d).



(a) Input Image (b)TSO (c)LBP (d) EBLD

Figure 3.13 (a) to (d) Unsupervised Texture Segmentation of Brodatz Images using TSO, LBP and EBLD with PC Method

The overall classification results for unsupervised texture classification of images using PC and FPC method are shown in Tables 3.2 and 3.3

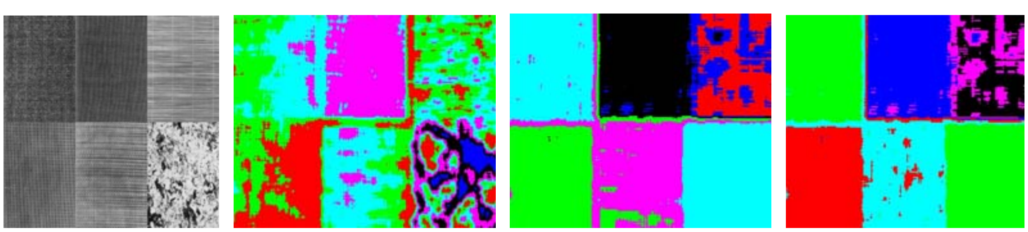
Table 3.2 Results for Unsupervised Texture classification of Brodatz Images of LBP and EBLD with PC method

Image	**Classification Percentage (%)**											
	Local Binary Pattern Operator with PC*						**Entropy Based Local Descriptor with PC***					
	D17	D21	D49	D53	D55	D60	D17	D21	D49	D53	D55	D60
D17	53.45	0	0.77	6.28	39.50	0	96.37	0	0.64	0	0	2.99
D21	0	81.04	4.56	0	13.16	1.24	0.16	93.99	0	5.33	0.52	0
D49	0	32.62	50.02	13.80	3.56	0	0.51	0	82.48	0	1.79	15.22
D53	0	1.07	0	62.08	0	36.85	8.80	0	2.40	88.77	0	0.03
D55	0	6.95	7.14	29.31	56.59	0	0	9.59	6.4	0	84.01	0
D60	24.92	14.30	11.92	9.44	14.05	25.35	1.56	0.34	3.85	0	0	94.25

* Proposed

From the Table, it is clearly seen that the proposed EBLD approach with PC has yielded better results with the overall classification accuracy of 96%.

The output images after applying unsupervised Fuzzy Partition Clustering classification methods in TSO, LBP and EBLD are shown in Figure 3.14 (b), (c) and (d).



(a)Input Image (b) TSO (c) LBP (d)EBLD

Figure 3.14 (a) to (d) Unsupervised Texture Segmentation of Brodatz Images using TSO,LBP and EBLD with FPC Method

Table 3.3 Results for Unsupervised Texture Classification of Brodatz Images of LBP and EBLD with Fuzzy Partition Clustering (FPC) Method

Image	Classification Percentage (%)											
	Local Binary Pattern Operator with FPC*						Entropy Based Local Descriptor with FPC*					
	D17	D21	D49	D53	D55	D60	D17	D21	D49	D53	D55	D60
D17	53.72	0	0.87	6.42	38.99	0	93.85	0	0.72	0	0	5.42
D21	0	81.18	4.51	0	13.04	1.27	0.09	92.26	0	5.26	2.39	0
D49	0	32.80	49.48	14.08	3.63	0.01	0.44	0	52.24	1.85	9.08	36.39
D53	0	1.09	0	62.94	0	35.97	10.46	0	0.47	89.00	0	0.07
D55	0	7.06	7.39	28.89	56.66	0	0	11.16	2.50	0	86.34	0
D60	24.39	13.72	12.10	10.37	14.37	25.03	0	0.40	1.97	0	0	97.63

*Proposed

From the Table 3.3, it is clearly evident that the proposed EBLD approach with PCM has yielded with the overall classification accuracy of 93%.

3.7 SUMMARY OF CONTRIBUTIONS

This chapter presents a statistical approach for texture description based on the classification methods of LBP, TSO and EBLD method. In the reason for misclassification in TSO method is due to overlapping the regions and the influence on the tolerance part. The LBP operator characterizes the spatial configuration of local image texture and its occurrence histogram is proven to be a very powerful texture feature results with the better classification accuracy than TSO. The EBLD classification approach is found useful for the images, which are having similarity in their shape where the features are extracted and successfully implemented for unsupervised classification. Among the three supervised methods EBLD maintains a high classification accuracy ratio with an average correct classification up to 96% and the percentage of correct classification has been tabulated.

The uses of different classification methods like LBP and EBLD Methods have been well described and the comparisons with the existing methods are also discussed in this Chapter. This EBLD method is tested with Brodatz images and performs well under different Gaussian noise levels with an average correct classification up to 96% has been obtained. The experiments show that the proposed LBP and EBLD methods can achieve better results than TSO texture classification approach. Hence, it is also found that the EBLD method with the combination of PC and FPC methods achieves high classification accuracy.

To improve classification accuracy further more using unsupervised hybrid methods are discussed in Chapter 4.

CHAPTER 4
IMPROVING CLASSIFICATION ACCURACY USING UNSUPERVISED HYBRID METHODS

4.1 INTRODUCTION

In the previous Chapter, the concepts of supervised and unsupervised classification using texture features are discussed. The unsupervised statistical methods have been very much used for discriminating different textures that includes Texture Spectrum Operator (TSO), Entropy Based Local Descriptor (EBLD), Local Binary Pattern Operator (LBP) and Gray Level Co-occurrence matrix (GLCM). Firstly the Texture Spectrum Operator (TSO) method has the merit that the texture aspects of an image are characterized by the corresponding texture spectrum instead of a set of texture measures and the texture spectrum can be directly used for image classification. On the other hand Local Binary Pattern Operator (LBP) method is even though suitable for many texture applications but with the implementation problem in the form of delta values definition form users to set the threshold values which makes it dependent on the gray scale values. For unsupervised texture classification analysis, the Gray Level Co-Occurrence Matrix (GLCM) method is suitable but with the shortcomings in classifying the large primitives. The Entropy Based Local Descriptor (EBLD) method has the disadvantage of considering the data in terms of probability resulting uncertainty and instability of data. In general, an image to be classified should undergo two basic processes namely feature extraction and determination of the optimal window size. Previously various methods, which have been used for classification in texture analysis by means of taking single image window size (fixed), created lesser classification accuracy. In order to improve the texture analysis classification quality, the advantages of existing different features could be integrated as hybrid techniques. This Chapter presents a new classification technique named 'Unsupervised Hybrid Classification for Texture Analysis' (UHCTA) by integrating different unsupervised methods to form a hybrid method that has the properties flexible to achieve higher classification rate by varying the window size for different images.

4.2 NEED OF HYBRID METHOD OF CLASSIFICATION

Texture classification is an image processing technique by which different regions of an image are identified based on texture properties. This process plays an important role in many industrial, biomedical and remote sensing applications. The statistical methods share one common weakness, of primarily focusing on the coupling between image pixels on a single scale and are also computationally intensive processes. Logical operators have been used for Boolean analysis, minimization, spectral layered network decomposition, spectral translation synthesis, image coding, cryptography and communication. Many texture classification methods already have been implemented and each one has its own advantages and disadvantages. In general any image which is to be classified should undergo two basic processes namely feature extraction and determination of the optimal window size. The existing methods, which have been used well for single image window size(fixed) that resulted some degree of compromise in achieving the classification accuracy and for better results in accuracy the advantages of existing different features could be taken and simultaneously to improve the classification accuracy.

For achieving better classification results in texture analysis, it is to combine different classification methods. Though there are existing methods which have been using fixed window size that resulted lack of classification accuracy and in order to improve the classification accuracy, the window size must be increased. Moreover the optimal window size selection is also an important thing for better classification output. In addition, some classification techniques are used for micro-textured structures and some are for large scale textured images and it is much needed to integrate different classification methods to achieve higher classification rate**.**

4.3 PROPOSED UNSUPERVISED HYBRID CLASSIFICATION FOR TEXTURE ANALYSIS METHOD (UHCTA)

This section focuses the implementation of the proposed Unsupervised Hybrid Classification Texture Analysis method (UHCTA). In this method, histogram techniques have been used for the unsupervised classification to compress the textured images. Once the images are selected from the Brodatz texture database in the proposed method, the local window is spilt into sub window of M×N, if window size is X. A sample of 20 images has been chosen for classification with the size of 512×512 for all the different unsupervised classification methods by varying different window sizes. The proposed Unsupervised Hybrid Classification Texture Analysis algorithm (UHCTA) is as follows.

Step 1: Initialize the process.
Step 2: Sample of different sets of images are taken.
Step 3: Compress the image dimension using histogram technique.
Step 4: Fix the optimal window size as 32 or as per the requirement.
Step 5: Window is spilt into 4 sub windows of 8×8, if window size is 32 which is optimal.
or Else no optimal
Step 6: Find classification rate.
Step 7: Finish procedure.

The proposed algorithm is explained in the flowchart as shown in Figure 4. 1.

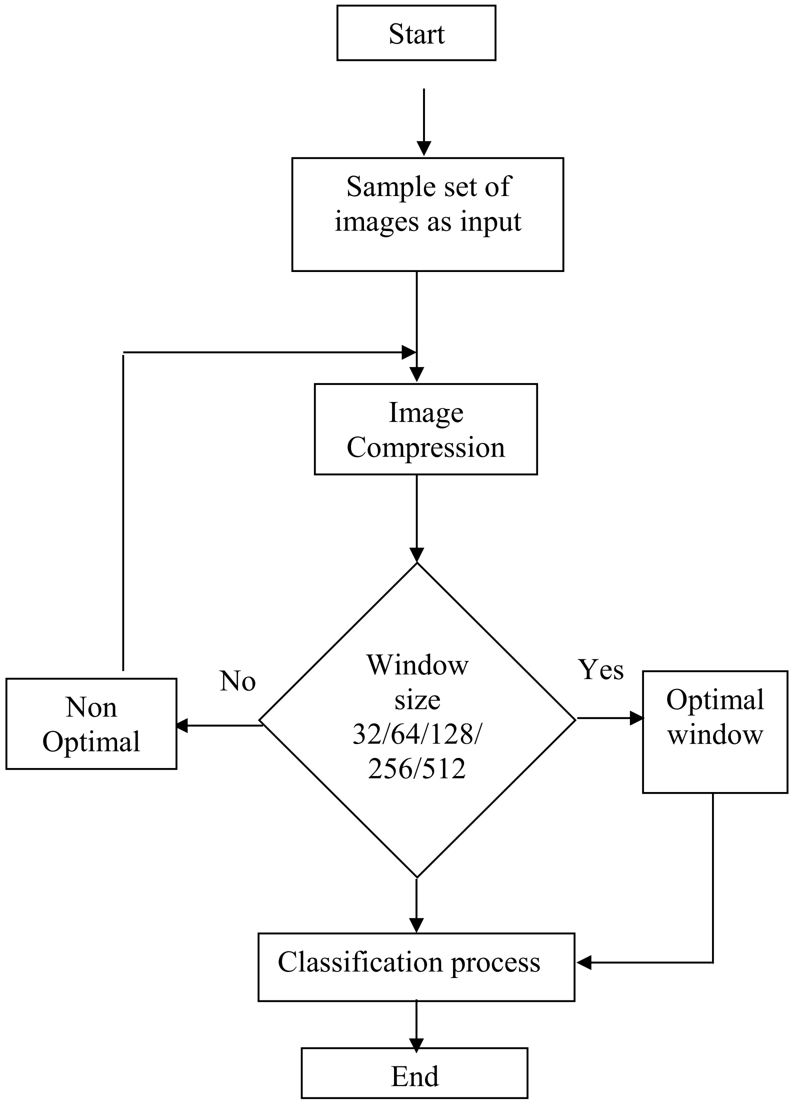


Figure 4.1 UHCTA Algorithm Flowchart

4.4 HISTOGRAM BASED METHODS

Histogram Based Methods are very efficient, when compared to other image segmentation methods because they typically require only one pass through the pixels. In this technique, a histogram is computed from all of the pixels in the image, and the peaks and valleys in the histogram are used to locate the clusters in the image. Color or intensity can be used as the measure. A refinement of this technique is to recursively apply the histogram-seeking method to clusters in the image in order to divide them into smaller clusters. This is repeated with smaller and smaller clusters until no more clusters are formed. One disadvantage of the histogram-seeking method is that it may be difficult to identify significant peaks and valleys in the image. In this technique of image classification, distance metric and integrated region matching are familiar.

Features of Histogram: Histogram-based approaches can also be quickly adapted to occur over multiple frames, while maintaining their single pass efficiency. The histogram can be done in multiple fashions when multiple frames are considered. The same approach that is taken with one frame can be applied to multiple, and after the results are merged, peaks and valleys that were previously difficult to identify are more likely to be distinguishable. The histogram can also be applied on a per pixel basis where the information results are used to determine the most frequent color for the pixel location. The histogram technique is not only for the compression but also reducing the computational cost.

Histogram Evaluation: An image often has defects in its contrast or brightness. It could be too dark, too bright and difficult to see in normal vision. It could be too rich in color that is not suitable for viewing. Histogram transformation is a technique that improves contrast by stretching the range of gray values to a desired level.

Histogram Equalization: This technique is commonly used in comparing images and for enhancing X-ray images. Apply a transformation function 'T_o' gray level 'u' to give gray level 'v' which has a uniform probability density.

$$V_k = \sum_{i=0}^{k} P_u(U_i) \qquad (4.1)$$

where, k is the gray level of the original pixel, and 'P_u' as the probability density of gray level u. In order to produce the transformation function to transform the pixels to the desired equalized level, Histogram normalization is required.

Histogram Normalization: This technique is commonly used in enhancing contrast of an image, while not distorting relative gray level intensities too significantly as histogram equalization would. Histogram normalization could simply be described as contrast stretching that is to stretch narrow histograms to a range specified to a user. First, we need to specify the upper and lower gray value limit in which, the output image should have, after normalization. The simplest type normalization takes in a pixel value and scales the values using this formula.

$$P_{out} = (P_{in} - c)\ (b-a/d-c) + a \qquad (4.2)$$

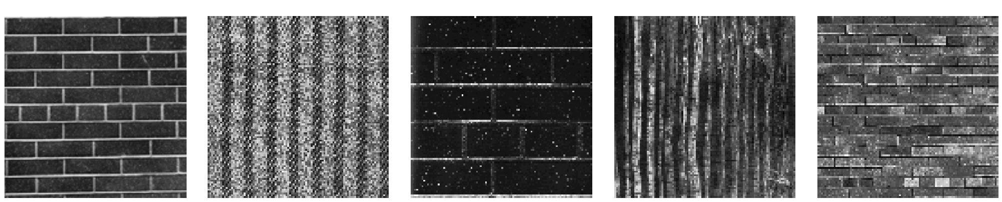
where Pin and Pout are the input and output gray value of the pixel respectively, with values below zero being set to zero and values greater than'd' being set to 'd'.

Local Spectral Histogram Representation: The spectral histogram provides a normalized feature statistic to compare image windows of different sizes. The input image windows do not need to be aligned; misalignment is a serious problem for approaches that use filter responses directly as features, such as those studied, due to the in homogeneity of filter responses. When proper filters are chosen, the spectral histogram is sufficient in characterizing texture appearance. Note that the spectral histogram is defined on any type of images. Piecewise-constant images with additive Gaussian noise are a special case where the spectral histogram has a unique pattern.

4.5 SIMULATION AND RESULTS

Here the proposed unsupervised hybrid classification method for texture analysis (UHCTA) method is tested to evaluate by choosing different window sizes respectively. The results of classification accuracy have been computed and compared with the different texture images shown in Table 4.1. The classification accuracy for Local Binary Pattern operator and Gray Level Co occurrence Matrix achieve less accuracy as compared with the proposed UHCTA. The sample images have been taken from the Brodatz data set.

The Brodatz images have been taken for the comparative analysis for both the cases of fixed window size and different window sizes as shown in Figure 4.2.



D26 Image D11 Image D25 Image D50 Image D9 Image

Figure 4.2 Brodatz Images

The comparative analysis for both the cases of fixed and different window sizes also has been done. At first, the proposed UHCTA algorithm is compared with LBP and it is found that the proposed classification method is so attractive in classification accuracy than the LBP while using fixed window size. The Figure 4.3 shows the average window size for UHTCTA is 90.6% which is higher than the 84 % of LBP.

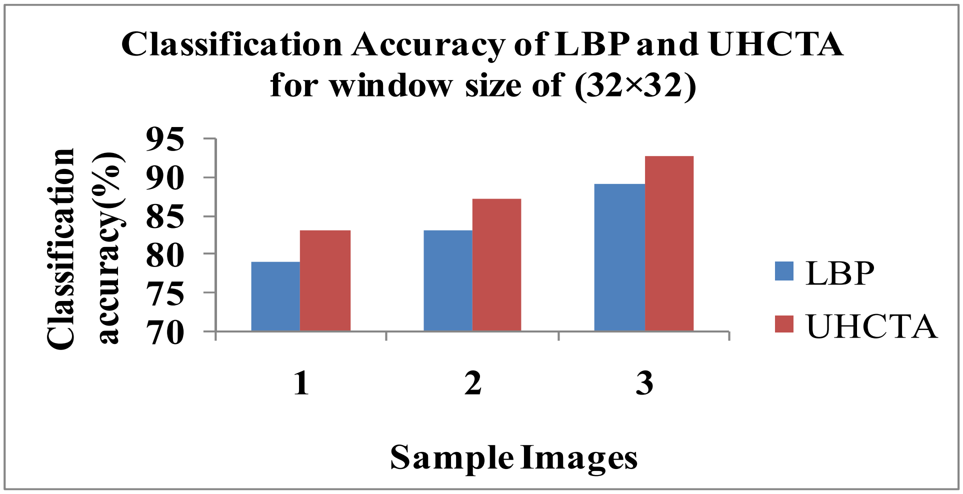


Figure 4.3 Classification Accuracy of LBP and UHCTA for window size of (32×32)

It is found that the proposed method has more classification accuracy than the other methods while using fixed window size. The average window size for UHTCTA is 90.6% which is higher than the 79.6% and 84% of LBP, and GLCM respectively as shown in Table 4.1.

Table 4.1 Average Classification Accuracy of LBP, GLCM and UHCTA for a Window Size of (32×32)

Classification methods	Window size of (32×32)			Average Percentage of Classification
LBP	79.00	83.00	89.00	79.00
GLCM	80.6	85.00	91.00	80.6
UHCTA	83.00	87.00	92.5	83.00

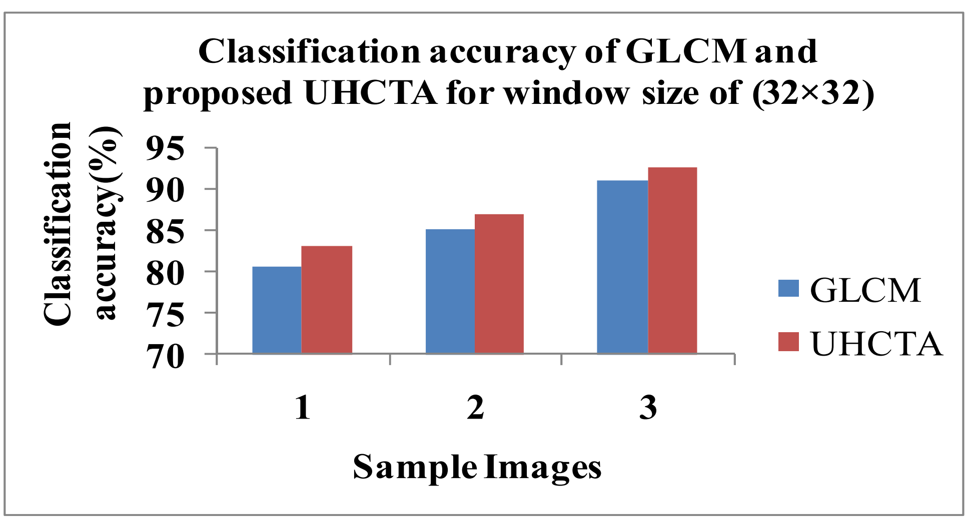


Figure 4.4 Classification Accuracy of GLCM and Proposed UHCTA for Window Size of (32×32)

Secondly the proposed UHCTA algorithm is compared with GLCM as shown in Figure 4.4.

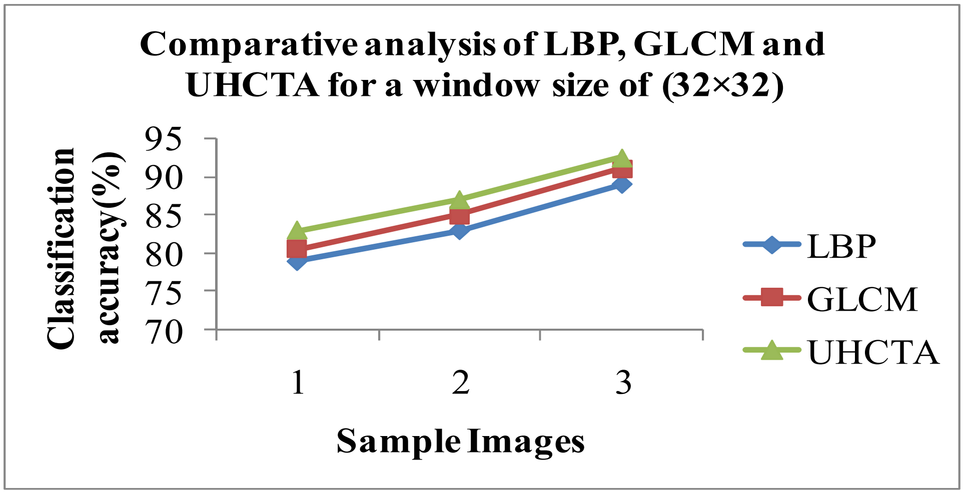


Figure 4.5 Comparative Analysis of LBP, GLCM and UHCTA for a Window Size of (32×32)

Table 4.2 Average Classification Accuracy of LBP, GLCM and UHCTA for a Window Size of (64×64)

Different classification methods	Optimal window size of (64×64)			
	Trial 1	Trial 2	Trial 3	Trial 4
LBP	83.5	85.00	87.5	88.00
GLCM	85.00	86.00	88.5	89.2
UHCTA	87.00	88.00	90.00	92.00

Table 4.2 shows the average classification accuracy of the LBP, GLCM and UHCTA methods and it is identified that the proposed UHCTA method outperforms the existing classification methods in terms of classification accuracy for the window size of 64×64.

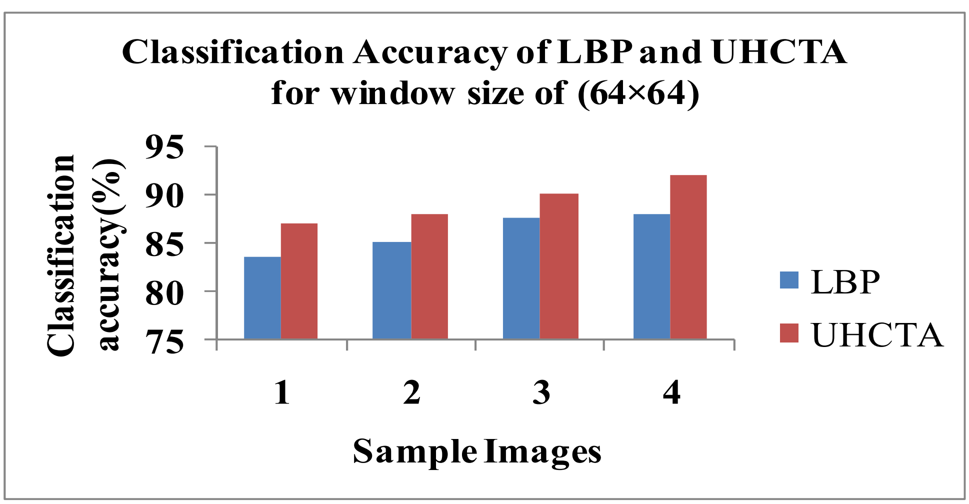


Figure 4.6 Classification Accuracy of LBP and UHCTA for Window Size of (64×64)

Figure 4.5 and 4.6 illustrates the comparative analysis of LBP, GLCM and the proposed UHCTA for a window size of 32×32 and 64×64 and it is concluded that the UHCTA method outperforms the others in achieving the classification accuracy.

The classification accuracy of GLCM is compared with the proposed UHCTA for a window size of 64×64 and it is inferred that the UHCTA has got attractive amount of classification accuracy than the GLCM which is shown in Figure 4.7.

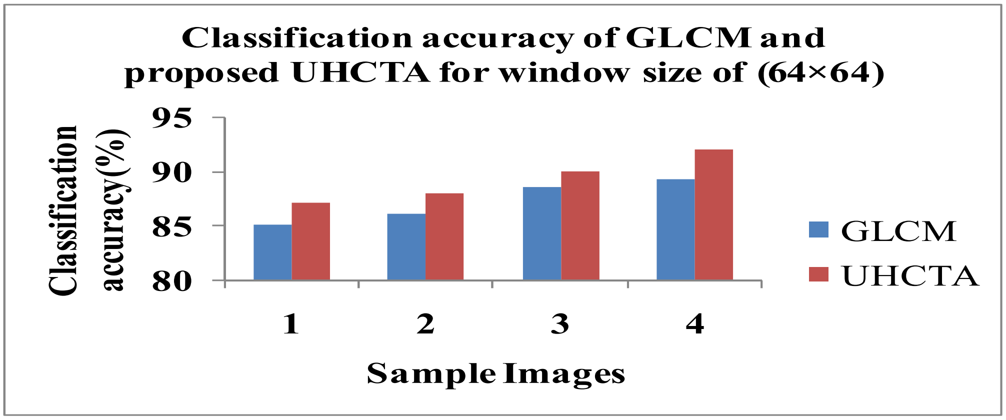


Figure 4.7 Classification Accuracy of GLCM and Proposed UHCTA for Window Size of (64×64)

Figure 4.8 shows the average classification accuracy of the LBP, GLCM and UHCTA methods and it is found that the proposed UHCTA method is so attractive than the existing classification methods in terms of classification accuracy for the window size of 64×64.

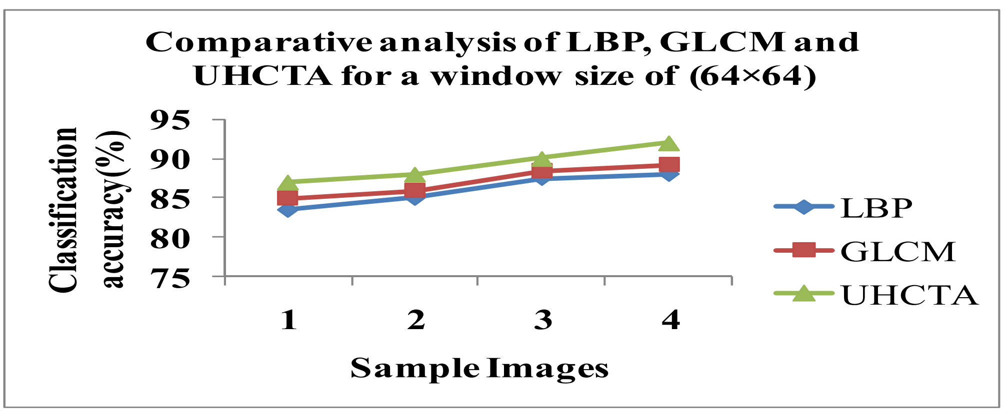


Figure 4.8 Comparative analysis of LBP, GLCM and UHCTA for a window size of (64×64)

Table 4.3 Average Classification Accuracy of LBP, GLCM and UHCTA for a window size of (128×128)

Different classification methods	Optimal window size of (128×128)			
	Trial 1	Trial 2	Trial 3	Trial 4
LBP	84.00	85.00	86.00	88.5
GLCM	85.00	86.00	88.00	89.5
UHCTA	87.00	88.5	90.00	92.5

Table 4.3 shows the average classification accuracy of the LBP, GLCM and UHCTA methods and it is identified that the proposed UHCTA method outperforms the existing classification methods in terms of classification accuracy for the window size of 128×128.

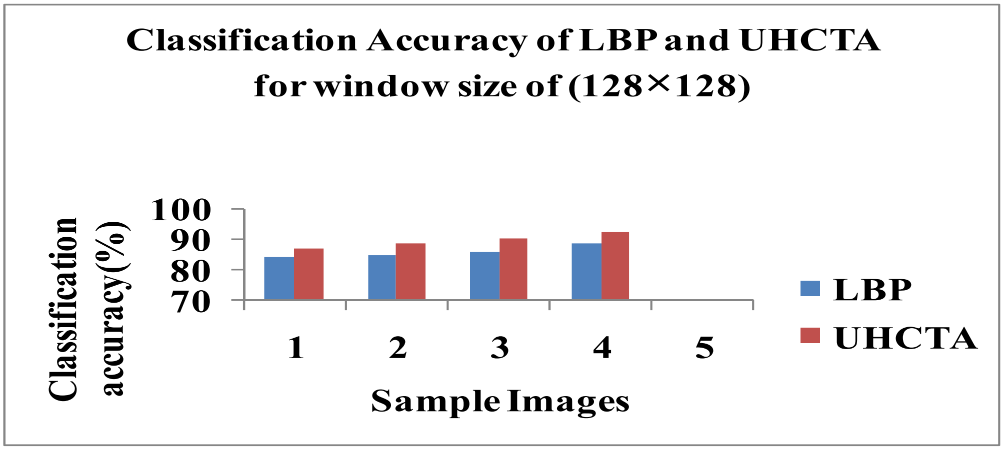


Figure 4.9 Classification Accuracy of LBP and UHCTA for a Window size of (128×128)

Figures 4.8 to 4.11 show the average classification accuracy of the GLCM and UHCTA methods and it is found that the proposed UHCTA method is so attractive than the existing classification methods in terms of classification accuracy for the window size of 128×128.

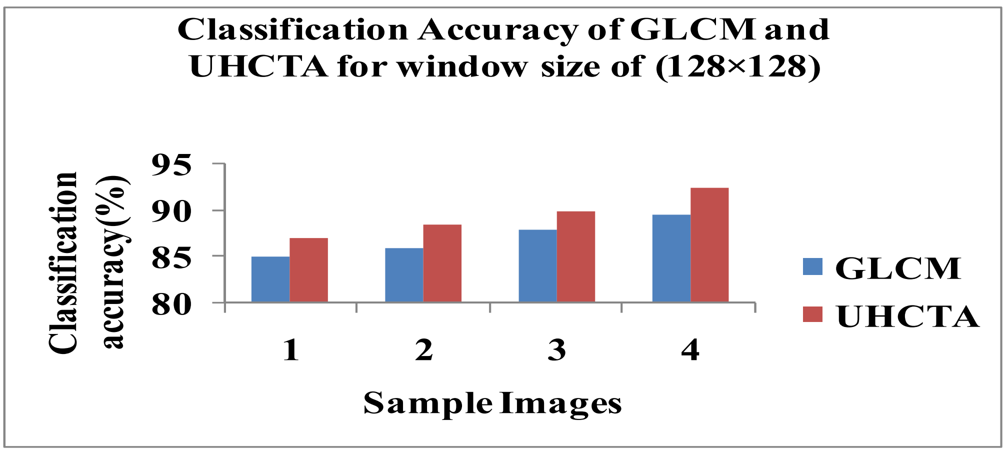


Figure 4.10 Classification Accuracy of GLCM and UHCTA for a Window Size of (128×128)

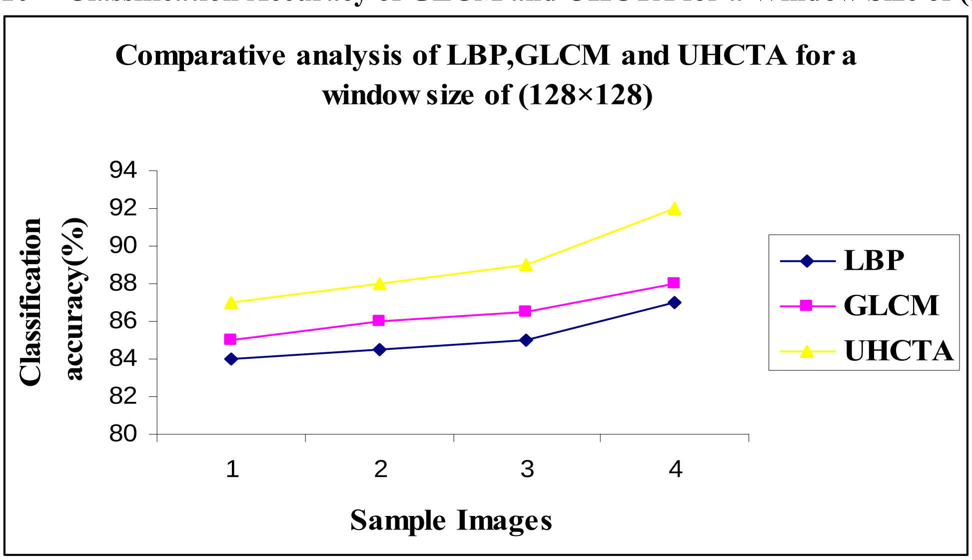


Figure 4.11 Comparative Analysis of LBP, GLCM and UHCTA for a Window Size of (128×128)

Table 4.4 Average Classification Accuracy of LBP and UHCTA for a Window Size of (256×256)

Different classification methods	**Optimal window size of (256×256)**			
	Trial 1	Trial 2	Trial 3	Trial 4
LBP	85.00	86.00	87.00	88.81
GLCM	87.00	88.00	89.00	90.00
UHCTA	88.5	89.5	90.5	92.00

Table 4.4 shows the average classification accuracy of the LBP, GLCM and UHCTA methods and it is identified that the proposed UHCTA method is far better than the existing classification methods in terms of classification accuracy for the window size of 256×256 and outputs are shown in Figures 4.11 to 4.14.

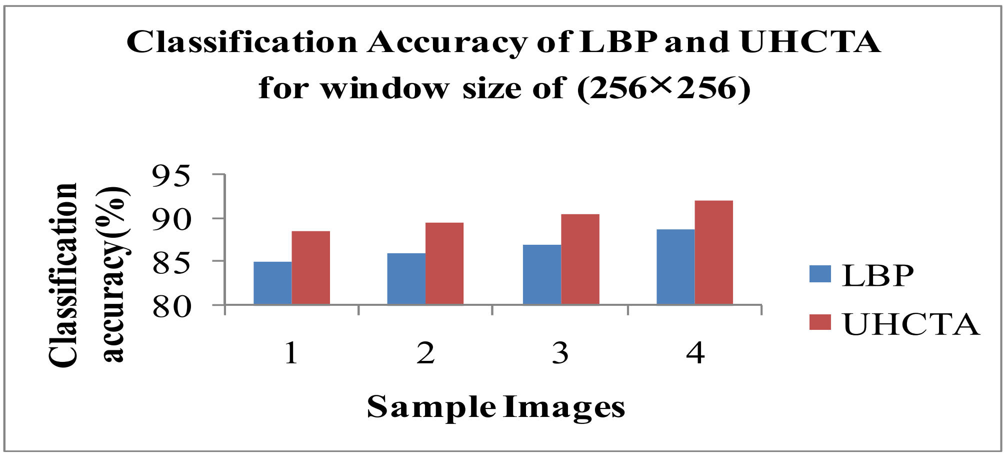


Figure 4.12 Classification Accuracy of LBP and UHCTA for a Window Size of (256×256)

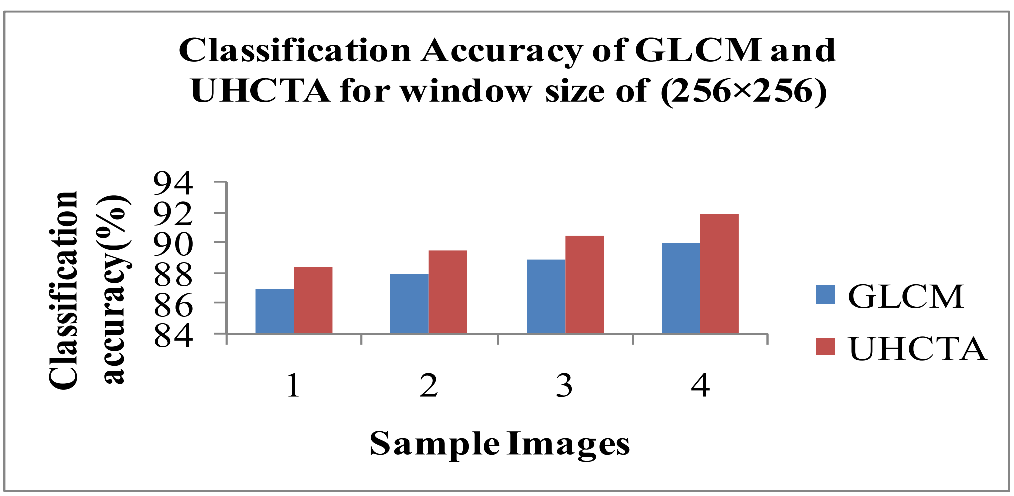


Figure 4.13 Classification Accuracy of GLCM and UHCTA for a Window Size of (256×256)

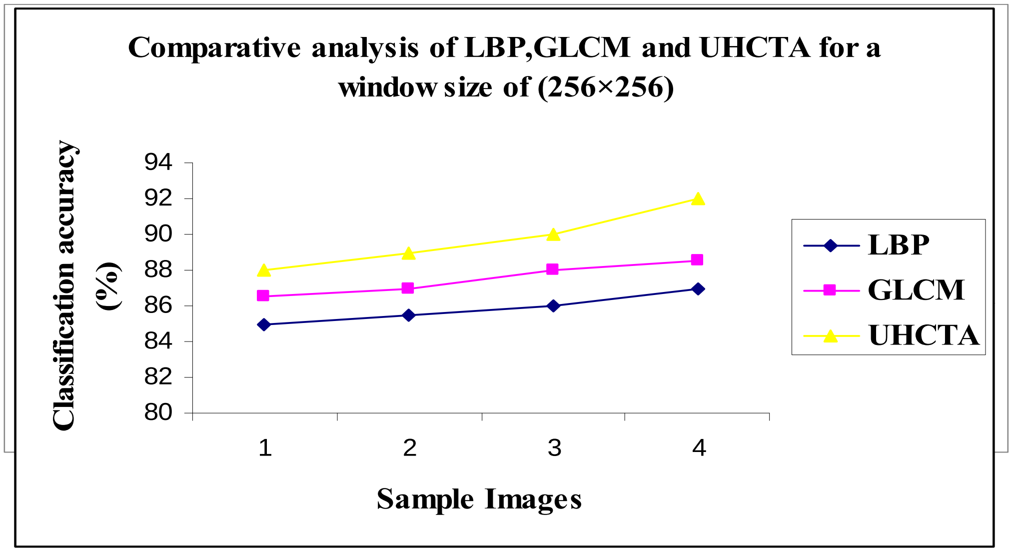


Figure 4.14 Comparative Analysis of LBP, GLCM and UHCTA for a Window Size of (256×256)

Table 4.5 Average Classification Accuracy of LBP, GLCM and UHCTA for a Window Size of (512×512)

Different classification methods	**Optimal window size of (512×512)**			
	Trial 1	**Trial 2**	**Trial 3**	**Trial 4**
LBP	86.00	87.00	87.5	89.00
GLCM	86.6	88.00	89.5	90.00
UHCTA	89.00	90.00	91.00	92.00

Table 4.5 shows the average classification accuracy of the LBP, GLCM and UHCTA methods and it is found that the proposed UHCTA method outperforms the existing classification methods in terms of classification accuracy for the window size of (512×512) as shown in Figures 4.14 to 4.16 with the classification of 92%.

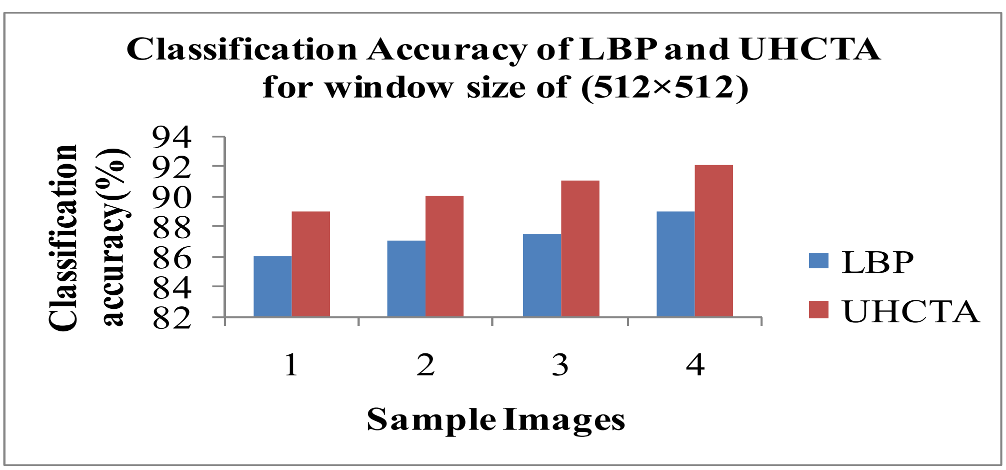


Figure 4.15 Classification Accuracy of LBP and UHCTA for a Window Size of (512×512)

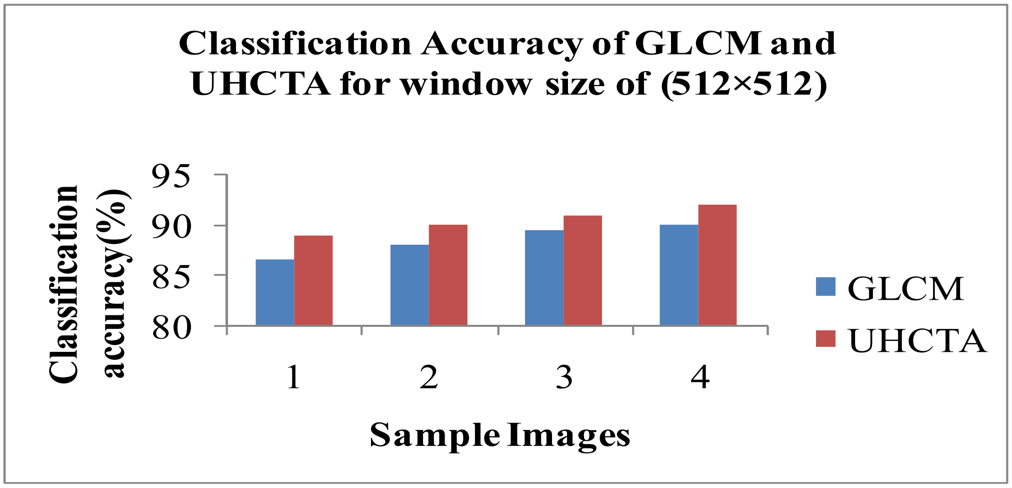


Figure 4.16 Classification Accuracy of GLCM and UHCTA for a Window Size of (512×512)

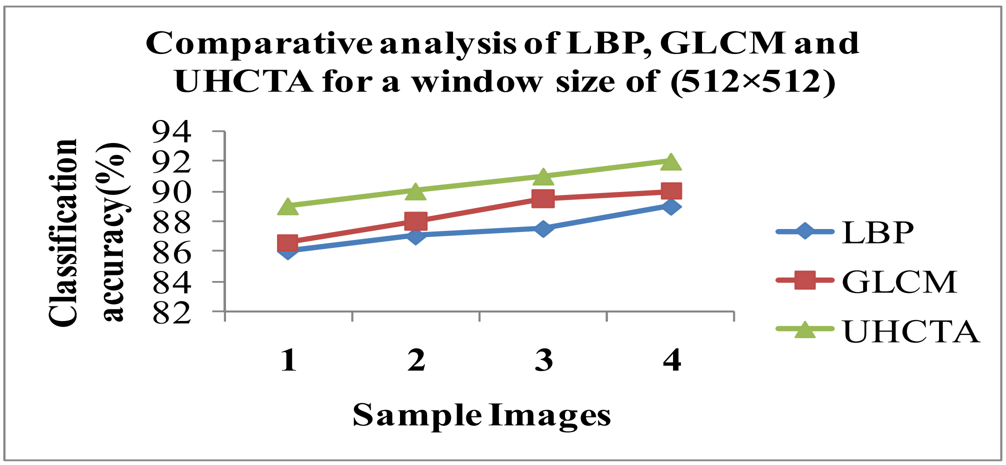


Figure 4.17 Comparative Analysis of LBP, GLCM and UHCTA for a Window Size of (512×512)

Figure 4.16 shows the average classification accuracy of the LBP, GLCM and UHCTA methods and it is found that the proposed UHCTA method works well than the existing classification methods in terms of classification accuracy for the window size of 512×512.

Finally the proposed UHCTA algorithm is compared with LBP, GLCM and UHCTA. It is found that the proposed method has attractive classification accuracy than the other methods while using different window sizes. The average window size for UHTCTA is 91.6% which is higher than the 85% and 86.6% of LBP, and GLCM respectively.

4.6 SUMMARY AND DISCUSSION

Unsupervised hybrid method in texture analysis has been implemented to improve the classification accuracy. The proposed method takes the advantages of various unsupervised texture analysis methods which is named as "hybrid" because of changing different textures by selecting optimal window size. The different image sets have been tested and experimental results conclude that the new method renders higher classification accuracy compared to previous methods. At first, the proposed UHCTA algorithm is compared with TSO and it is found that the proposed method is so attractive in classification accuracy than the TSO while using fixed window size. The average window size for UHTCTA is 90.6% which is higher than the 77% of TSO. Secondly the proposed

UHCTA algorithm is compared with EBLD, LBP and UHCTA and it is found that the proposed method has more classification accuracy than the other methods while using fixed window size. The average window size for UHCTA is 90.6% which is higher than that of 79.6% and 84% of EBLD, and LBP respectively. Finally the proposed UHCTA algorithm is compared with LBP, GLCM and UHCTA while using different window sizes and it is found that the proposed method outperforms the other methods in classification accuracy. The average window size for UHCTA is 91.6% which is higher than the 85% and 86.6% of LBP, and GLCM respectively for the different window size.

This chapter presents a design of a Hybrid unsupervised classification method to increase the classification accuracy which provides higher degree of classification rate. In addition by varying window size, the classification accuracy has been increased. An unsupervised hybrid classification for texture analysis has been implemented that comprises the features of various classification methods to improve the classification accuracy. The proposed technique contains three issue that are selection of window size, fixing the optimal size of the window and compressing the texture image by means of splitting into subsets. After applying hybrid features, different sets of images have been tested. From the experimental results, it is concluded that the proposed method imparts higher classification rate than the existing method with the average classification accuracy of 92%, which is more attractive.

The design of a pixel based classification on color images in statistical texture analysis and comparative analysis are proposed in **Chapter 5.**

CHAPTER 5
PIXEL BASED CLASSIFICATION ON COLOR IMAGES IN STATISTICAL TEXTURE ANALYSIS

5.1 INTRODUCTION

Design of an Unsupervised Hybrid Classification for Texture Analysis (UHCTA) algorithm by varying the window size for different images has been discussed in the previous chapter. When using statistical approach in texture analysis for image, classification (Particularly Gray Level Co-occurrence Matrices approach) is applied in discriminating different textures in images results better accuracy but with the high computational cost. In addition, a single window classification for each pixel compromises the classification accuracy rate. This chapter concentrates on image classification implementation by using a statistical method in pixel by pixel with maximum likelihood estimates. The experiments are to be tested with the color images to improve the classification rates and then to be compared with a single window classification. Previously a number of different texture analysis methods have been introduced namely statistical, structural, transform based and model based methods (Castielloc et al 2003). The statistical method has the main features which are to be extracted that includes the autocorrelation function, power spectra, difference gray level statistics, co-occurrence matrices and from sum and different statistics (Argenti et al 1990). On the other hand a new method is to be introduced to classify color images based on the study of statistics of the color components.

In this chapter, design of a pixel based classification on color images in statistical texture analysis and comparative analysis of various images are proposed.

5.2 PROBLEM DEFINITION

Previously many authors have discussed on the image classification in statistical texture analysis based on pixel and on color images. Argenti et al (1990) proposed a fast algorithm for texture analysis using co-occurrence matrices to improve the classification accuracy. Acha et al (2000) presented a method to classify color images in to different groups based on texture and color images. Jaime Melenbez, 2010, described a novel method to improve classification accuracy of the common co-occurrence matrix approach on standard textures significantly. Bhowmick et al (2009) combined co-occurrence matrices and the scale space approach for texture analysis by regarding the distance‘d’ as scale parameter. Fernando Bello, 1996 compared filtering approaches with co-occurrence matrices and found some filters performing superior for selected textures. But no authors so far proposed a method for classifying color images based on pixel by pixel using co-occurrences matrices. So this chapter aims at achieving a better classification accuracy on color images, which is based on statistical method.

The main objective of this chapter is to classify the images on a pixel by pixel basis, where each pixel is associated with textural features extracted from co-occurrence matrices that differs the pixel itself.

5.3 COLOR SPACES AND TEXTURE CLASSIFICATION

The color of pixels can be represented in different color spaces which respect different physical, physiologic, and psycho-visual properties. They can be classified into four families: the primary spaces, the luminance chrominance spaces, the perceptual spaces and the independent color component spaces.

5.3.1 Color Texture Features

For the texture classification purposes, a color texture is described by a set of ‘d’ features and is represented by a point in a d-dimensional feature space in order to achieve the classification. There exist a large number of color texture descriptors and it is well-known that the performance of the classifier depends on the choice of the color texture features. In order to represent both the color and the texture information, three kinds of features are used in the literature:

- “Luminance Based Texture Features” mixed with color statistical features.
- “Within color component texture features” which takes into account only the spatial relationships within a single color component (for example within the color component R, G or B).
- “Between color component texture features” which consider spatial relationships within and between different color components.

5.3.2 Feature Selection on Color Images

The supervised classification scheme for color image classification is divided into two successive stages:

- A supervised procedure selects a low number of discriminating texture features among a set of candidate ones in order to build a low-dimensional feature space. During this stage, the classifier is trained to partition this feature space.

- A decision stage where each examined color texture is represented by a point in the selected feature space in order to be classified.

5.4 COLOR TEXTURE ANALYSIS METHODS AND THEIR NEED

Texture features characterize the statistical or structural relationship between pixels, and provide measures of properties such as contrast, smoothness, coarseness, randomness, regularity, linearity, directionality, periodicity, and structural complexity. Image clustering or classification is one crucial step for image analysis. Color image consists of three layers Red, Green and Blue (RGB). The color composition could be represented in a RGB space that shows any color is a combination of red, green and blue elements. A basic color image could be described as three-layered image, with each layer as red, green and blue. To consider a particular area of a color image, it could have different color combinations with respect to other areas of the image. An area with yellow would have a combination of red and green color elements only. Similarly, an area with magenta would have combination of red and blue elements only. Thus, the object is best described in red and green layers, not in blue layer.

Texture classification assigns a given texture to some texture classes. Two main classification methods are supervised and unsupervised classification. Supervised classification is provided examples of each texture class as a training set. A supervised classifier is trained using the set to learn a characterization for each texture class. Unsupervised classification does not require prior knowledge, which is able to automatically discover different classes from input textures. Another class is semi-supervised with only partial prior knowledge being available. So color texture analysis is always very difficult method and therefore a better classification method is highly required.

5.5 PROPOSED PIXEL BY PIXEL BASED IMAGE CLASSIFICATION ON COLOR IMAGES

Any image that comprises small entities or cells, whose shape, dimensions and spatial distribution distinguishes different texture classes. To classify the texture, those entities sizes are to be compared to pixel resolution. In order to find the spatial relationships effectively, the classification method is used and Grey-Level Co-occurrence Matrix (GLCM) is one of the most widely used statistical texture measures. The idea of the method is to consider the relative frequencies for which two neighboring pixels are separated by a distance on the image. Since the GLCM collects information about pixel pairs instead of single pixels and which is called by a name as second-order statistics. Texture measures, such as homogeneity, contrast, and entropy are derived from the co-occurrence matrix. The different sets of color images of vistex dataset have been tested for the classifications. For color texture characterization, the essential statistical parameters are Kurtosis and skewness, normalized moments of third and fourth order respectively.

The color image is represented by three co-ordinates: the luminance (L) and the two chrominance C1 and C2. To characterize the images by the centroid vector of the two chrominance (C1, C2), each one represents a pixel. The proposed algorithm for the classification of color images using co-occurrences matrices approach is explained as follows.

A pixel neighborhood can be characterized by color texture feature values which are computed by tacking into account the color components of the neighbor player pixels. Here we use a non exhaustive list of texture features. The mean of the pixel values in a neighborhood, the median and mode evaluate the central value of this neighborhood. The variability of the pixel values around a central value is estimated by the variance or its square root. The skewness estimates the degree of asymmetry of the pixel values around a central value. The variance and the skewness can be evaluated around the mean, the median or the mode. Let N_t, be the number of available texture features, which are computed with one color component. The pixels of a color image are usually digitized with the (R, G, B) color representation system. Nevertheless the R, G and B color components are not always adapted to a specific problem of color image segmentation. In digital color imaging, many other color representation systems exist. So, the objective is to look at the best subset of color texture features for discriminating the different classes of pixels.

5.5.1 Color Texture Feature Space Selection

In order to reduce the dimension of color texture feature space selection, we use an iterative feature selection procedure. At each step k of this procedure, we consider several candidate of color texture feature spaces for which we compute their discriminating power thanks to an information criterion J. At the beginning of this procedure (k=1), we consider the Nf mono-dimensional candidate space which maximizes J is the best one for the N classes. We select this space as the first one and we associate it, in the second step of the procedure (k=2), to each of the (Nf-1) remaining candidate color texture features in order to constitute (Nf-1) bi-dimensional candidate spaces. We consider that the bi-dimensional space which maximizes J is the best plane for discriminating the classes. The procedure is iterated until stabilization of the value of J. Let ko, the rank of the iteration process, which corresponds to the beginning of the stabilization of J. ko is the dimension D of the color texture feature space. This classical multiple discriminant analysis method does not yield the optimal solution but a satisfying one which is less computation time consuming.

The evaluation of the discriminating power supposes that the more the classes are well separated and compact in the candidate color texture feature space, the higher the discriminating power of the selected feature .The new color image segmentation approach proposed here shows the contribution of color texture features to supervised pixel classification. The determined color texture feature space is specific to an application and depends on the chosen texture features and the criterion used.

5.5.2 Object Based Feature Extraction Approaches

Object based feature extraction is a new method, which is widely used recently in huge number of studies in order to estimate more accurate results. Object based image analysis approach is the approach to image analysis combining spectral information and spatial information, so with object base approach not only the spectral information in the image will be used as classification information, the texture and context information in the image will be combined into classification as well (Flanders et. al 2003). The object based classification concept is that important semantic information necessary to interpret an image is not represented in single pixels, but in meaningful image objects and their mutual relations. Image analysis is based on contiguous, homogeneous image regions that are generated by initial image segmentation. Connecting all the regions, the image content is represented as a network of image objects. These image objects act as the building blocks for the subsequent image analysis. In comparison to pixels, image objects carry much more useful information.

5.6 SUPERVISED PIXEL BASED CLASSIFICATION

A pixel based classification aims at determining the class to which every pixel of an input image belongs based on several measures computed by applying a set of texture feature extraction method such as Gabor filters which leads to the segmentation of the image as a collateral effect. In order to take the advantage of the output of Gabor filter method texture features have been computed for different evaluation window sizes and processed during classification by following the top down approach described below.

1. Select the largest available evaluation window and classify the test image pixel labeled as unknown. (Initially all pixels are labeled as unknown).
2. Locate the pixels of the classified image that constitute the frontiers between regions of different texture and marked them as unknown as well as their neighborhoods. The size of the neighborhood corresponds to the size of the evaluation used to classify the images.
3. Discard the current evaluation window.
4. Repeat the steps 1 to 3 until the smallest window has been utilized.

In this way, large windows are applied inside regions of homogeneous texture in order to avoid noisy classified pixels and small windows are applied near the frontiers between those regions in order to refine them. Further more the above strategy renders classifying every image pixel with the entire available window unnecessary. Hence it leads to low computation times.

5.7 PIXEL BASED SEGMENTATION

The segmentation technique which operates in color space is broadly divided into three groups.

1. Histogram based techniques: one or more peaks are identified; surrounding intervals are utilized next in pixel classification process.

2. Segmentation by clustering data color spaces: pixel values are collected into groups with one or more representatives, which are used in pixel classification.

3. Segmentation by Fuzzy Clustering: Fuzzy membership functions are evaluated for all pixels and for all Fuzzy clusters defined, hard clusters of pixels are obtained by defuzzification process and subdivided into maximal connected regions.

5.8 COLOR IMAGE SEGMENTATION

Image segmentation that is identification of homogeneous regions in the image has been the difficult task. Many algorithms have been elaborated for gray scale images. However the problem of segmentation for color images, which convey much more information about objects in scenes. Color image segmentation is a process of extracting from the image domain one or more connected regions satisfying uniformity (homogeneity) criterion, which is based on features derive from spectral components. These components are defined in a chosen color space model. Perhaps the most important feature of the segmentation method is region definition. Region is a connected set of pixels for which uniformity condition is satisfied. For instance

a) Uniform region derived by growing from a seed block by joining other pixels or block of pixels.

b) Uniform region obtained by splitting a larger region which is not uniform. Region is a connected set of pixels bounded by edge pixels creating a color contour or region corresponds to a surface or an object of homogeneous materials.

5.9 PROPOSED PIXEL BASED CO-OCCURRENCE METHOD (PBCM) ALGORITHM

Supervised texture segmentation identifies and separates regions that match texture properties previously learnt in training samples. In turn, unsupervised texture segmentation can be supervised or unsupervised depending on whether prior knowledge about the image or its texture classes is available or not. The proposed PBCM method is explained in steps as follows.

Step 1: The first window W(x, y) relates to the pixel (k, l) of f(x, y) is fixed.
Step 2: The neighboring window $\acute{W}'(x, y)$ that relates to (k+1,l) is fixed.
Step 3: Pixels separated by δ in the neighboring window W(x ,y) and $\acute{W}'(x, y)$.
Step 4: Co-occurrence matrix $\acute{M}\ s^{1}(x, y)$ relates to $\acute{W}'(x, y)$ is got by updating the matrix relative to W(x, y).
Step 5: Decrementing by one entries $\acute{M}\ \delta\,(x, y)$ due to the pairs of the left hand columns and incrementing in the right hand column yields $\acute{M}\ s^{1}(x, y)$.
Step 6: The parameters extracted from co-occurrence matrices relates to $\acute{W}'(x, y)$ is calculated by updating the adjacent pixel (k, l).
Step 7: Co-occurrences parameters are to be found, where the mean and standard deviation of row sums of matrix and are analogous statistics of the column sums.

The flow chart for the above algorithm has been illustrated in Figure 5.1.

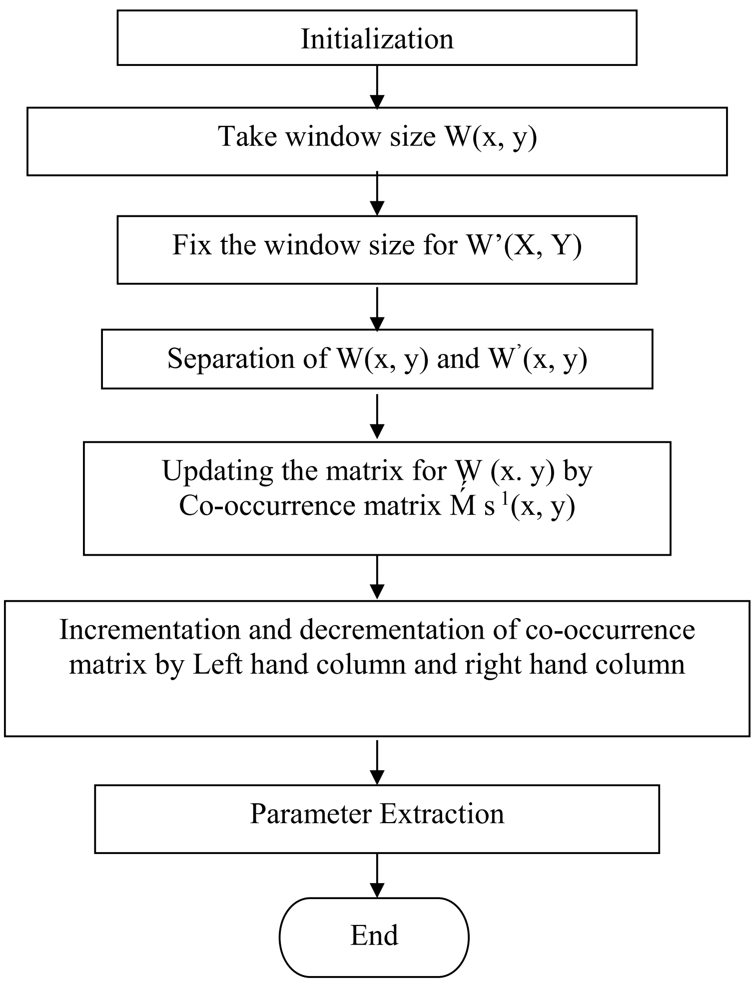
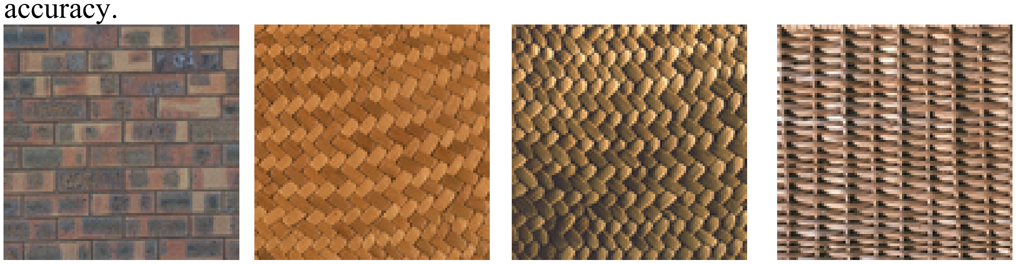


Figure 5.1 Co-occurrence Matrices for Parameter Extraction Algorithm

5.10 EXPERIMENTAL RESULTS AND OUTPUT

30 color images have been obtained from vistex dataset and the two statistical parameters kurtosis and skewness are taken for the classification purpose. The sample image from vistex (brick) is shown in Figure 5.2 (a) to (d). Three different window sizes have been taken for the experimental purpose and it is confirmed that there is a clear difference in the classification accuracy.



(a) (b) (c) (d)

Figure 5.2 (a) to (d) Sample Brick images from Vistex Data Base

Table 5.1 Comparison Between Proposed PBCM Algorithm and GLCM

Different classification methods	Different window size Classification Accuracy (%)		
	8×8	16×16	64×64
Existing Methods(GLCM)	83.00	85.00	87.00
Pixel Based Co-occurrence Method (PBCM)	86.00	88.00	90.00

The co-occurrence matrices parameters are extracted by the proposed algorithm and the experimental results are compared with the existing work (Argenti et al 1990) shown in the Table 5.1.

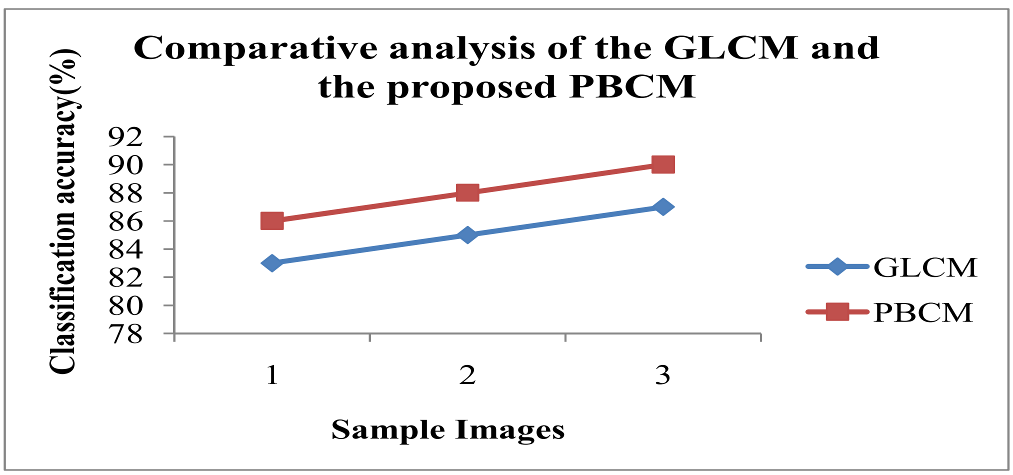


Figure 5.3 Comparative Analysis of the GLCM and the Proposed PBCM Method

Figure 5.3 shows the classification accuracy of the existing method and the proposed PBCM method in obtaining the classification accuracy. Figure 5.4 shows the classification accuracy of the existing method and the proposed method for the window size of 8×8 and it is noted that the proposed method is so attractive in obtaining the classification accuracy.

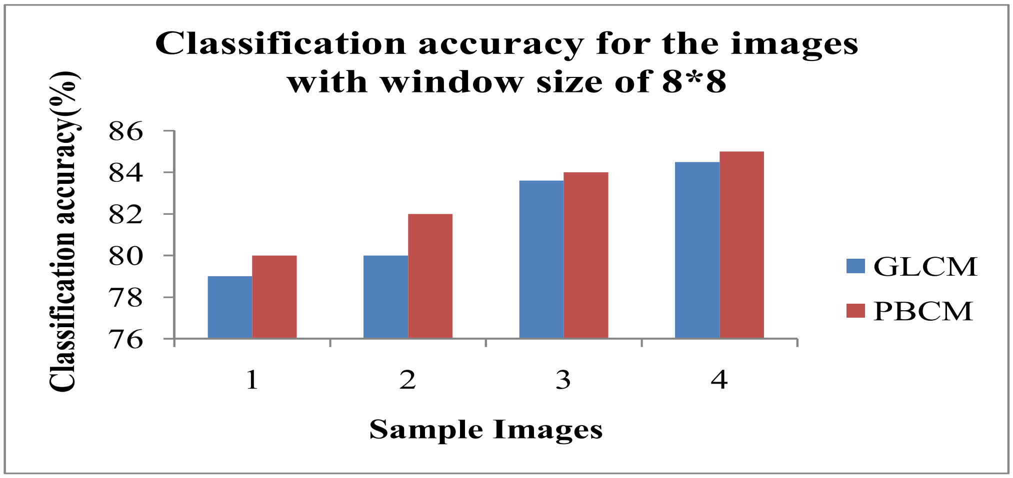


Figure 5.4 Classification Accuracy for the Images with Window Size of 8×8 of the GLCM and the Proposed PBCM Method

Figure 5.5 shows the classification accuracy of the GLCM method and the proposed method for the window size of 16×16 and it is found that the proposed method is so attractive in obtaining the classification accuracy.

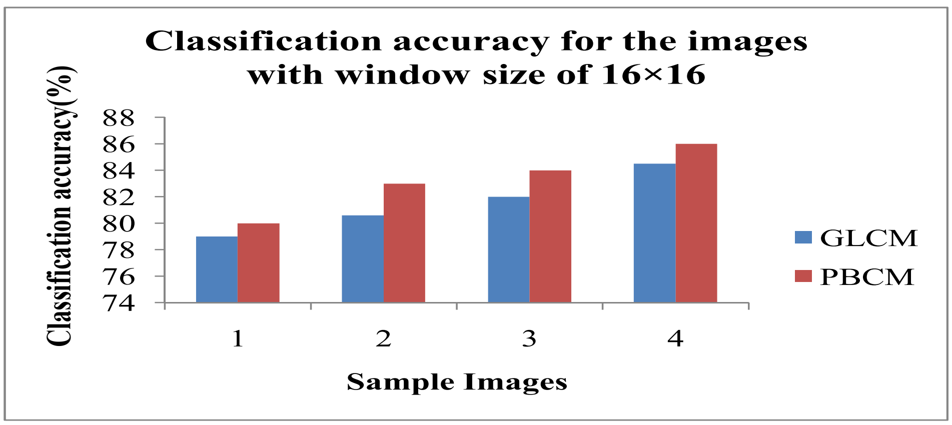


Figure 5.5 Classification accuracy for the images with a window size 16×16 of GLCM and the proposed PBCM method

Figure 5.6 shows the classification accuracy of the existing method and the proposed method for the window size of 52×52 and it is studied that the PBCM method outperforms in obtaining the classification accuracy.

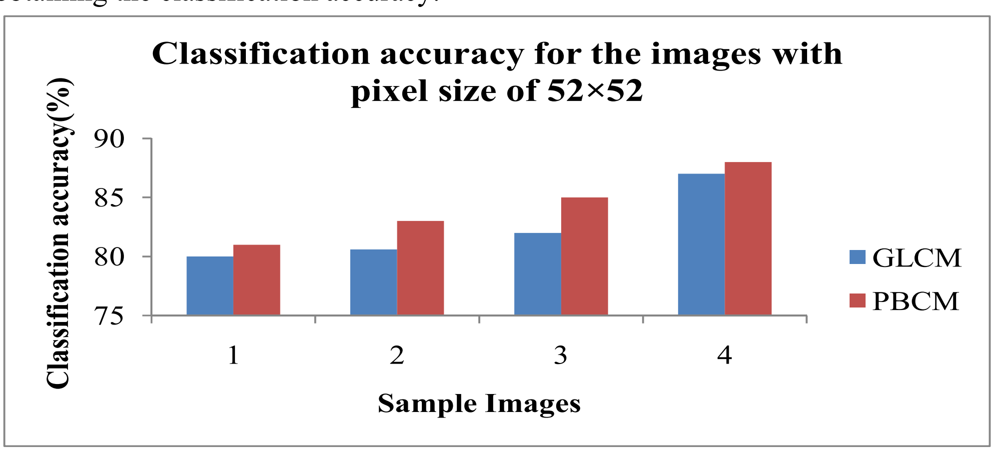


Figure 5.6 Classification accuracy for the images with a window size 52×52 for the GLCM and the proposed PBCM method

Figure 5.7 show the classification accuracy of the GLCM and the PBCM method for the window size of 64×64 and it is clearly studied that the proposed method has a high level of classification accuracy.

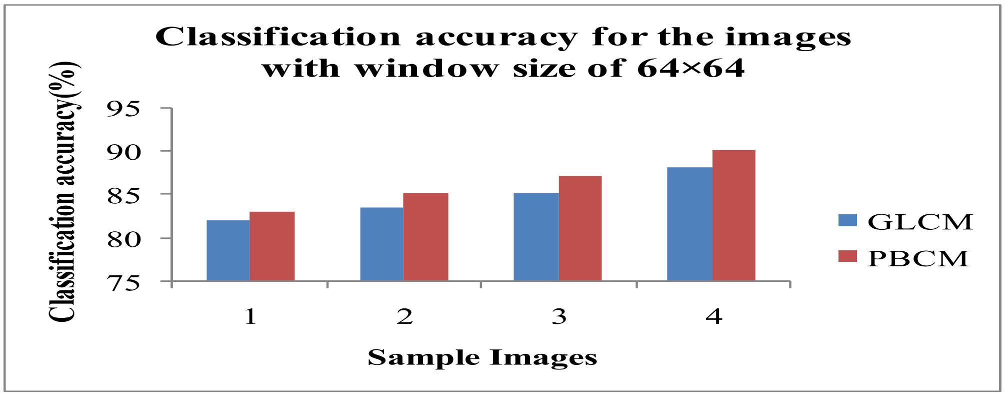


Figure 5.7 Classification Accuracy for the Images with Pixels 64×64 for GLCM and the Proposed PBCM Method

Figure 5.8 shows the classification accuracy of the GLCM and the proposed method for the window size of 126×126 and it is found that the PBCM method is so good in getting the classification accuracy.

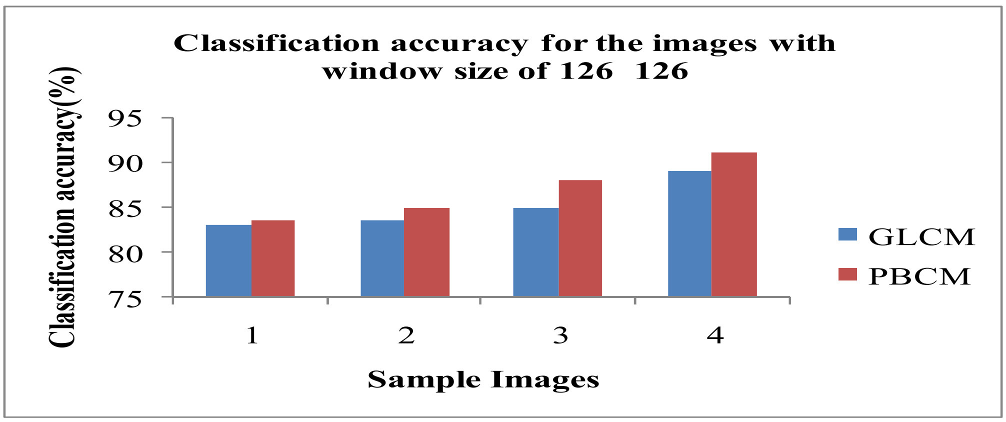


Figure 5.8 Classification Accuracy for the Images with Window Size 126×126 of GLCM and the Proposed PBCM Method

The classification accuracy of the proposed algorithm have been calculated after extracting the co-occurrences parameters and it is found that by applying pixel by pixel approach, there is a better classification accuracy as shown in Figure 5.9.

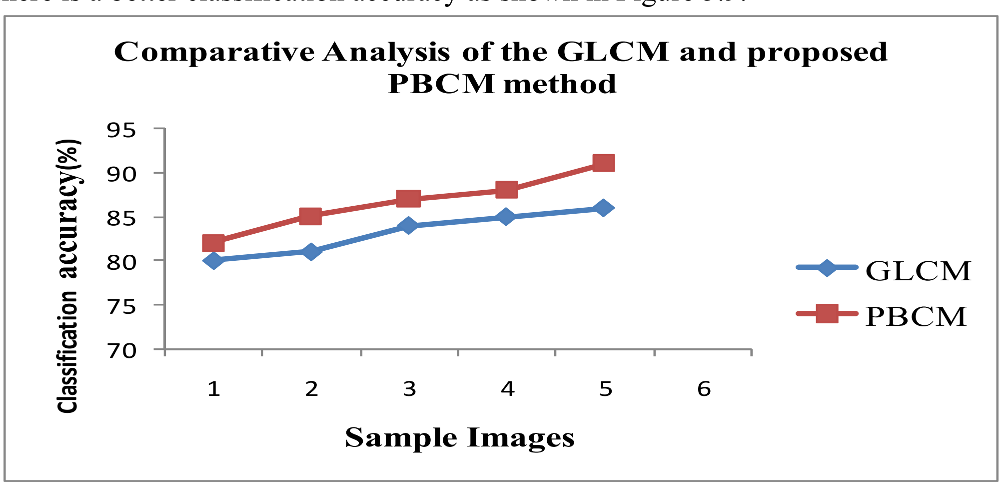


Figure 5.9: Comparative Analysis of the existing method GLCM and proposed PBCM method

The overall comparison has been manipulated and plotted in the Table 5.2 and it is clearly understood that the proposed PBCM method has high amount of classification accuracy comparing with the existing method for different window sizes.

Table 5.2 Classification accuracy of the existing method GLCM and the proposed PBCM method for the different window sizes

Different classification methods	Different window sizes [Classification Accuracy (%)]				
	8×8	16×16	52×52	64×64	126×126
GLCM	80.00	81.00	84.00	85.00	86.00
Pixel Based Co-occurrence Method (PBCM)	82.00	85.00	86	89.00	91.00

The classification accuracy of the proposed algorithm have been calculated after extracting the co-occurrence parameters and it is found that by applying pixel to pixel approach, there is a better classification accuracy.

5.11 CONCLUSION

The proposed classification algorithm based on pixel by pixel method on color images has been implemented and applied. The features on color images are extracted by means of GLCM statistical method through the concept of overlapping window for neighboring pixels. The experiments were done for different window sizes and the proposed PBCM method is also compared with the existing method and it is concluded that the classification accuracy of the color images have been in higher rate. Here the tests were conducted for the color images by overlapping neighbors for the supervised and unsupervised methods. The results of the proposed algorithm may be very useful in medical image applications, where more accuracy in terms of color image classification is required. Comparative analysis of the proposed with the existing is also done and the result shows that the classification accuracy is so attractive.

The proposed classification method on mammogram in detecting micro-calcification is explained in the next chapter.

CHAPTER 6
STATISTICAL TEXTURE ANALYSIS FOR MEDICAL IMAGES

6.1 INTRODUCTION

In the previous chapter, the texture classification on color images based on pixel by pixel method was studied. Texture analysis has been widely used in medical image tasks. Despite their success in image retrieval and category classification, bag-of-words models have not been thoroughly studied for image-based diagnosis tasks. In that, analysis of the mammographic texture, using various texture features is the most difficult one. Among all medical image task detection, micro-calcification on mammograms is the most difficult one, because breast cancer is the most prevalent cancer (Deepa Sankar 2009). More over micro-calcification are deposits of calcium that can be seen in mammograms and which is the best way to detect breast cancer in the earliest stage (Mascaro et al 2009). In the unsupervised classification method for the detection of micro-calcification, the prior information is not required and in the case of supervised method, information on micro-classification is very much needed. A number of techniques for detection of micro-calcifications have been developed in the past decade. Previously various methods have been developed for the detection of micro-calcification on mammogram and none of them have concluded the identification of best classification methods. The aim of this chapter is to study and conclude the suitability of different classification methods in detection of micro-calcification in breast cancer by applying common sets of images with the various supervised methods and unsupervised methods to detect the classification accuracy. In addition, the use of supervised and unsupervised classification methods in the way of detecting micro-calcification to achieve a good 'True Positive' detection and at the same time as low a 'False Positive' detection rate as is possible by using the same set of MIAS image analysis have been studied.

6.2 NEED FOR CLASSIFICATION OF MAMMOGRAM

An accurate computer aided diagnosis systems can be more helpful for radiologist to detect and diagnose micro-calcification percentage in advance than the normal screening test. The early detection of breast cancer is done by a special type of x-ray called mammography. The classification of calcium deposits into the suspicious and normal categories is a difficult task, which includes the detection of tumors as mass regions with a weak contrast to their background and the extraction of features which characterize malignant tumors. Owing to the small size of micro-calcification with a diameter of less than 0.5 mm level and are in the form of groups as clusters and in homogeneous back ground, it is very difficult to detect. Moreover the detection of masses in mammograms are too sensitive and so it is better to use the different supervised classification and unsupervised methods on mammographic images in order to describe the various discriminate characteristics of both true and false phenomenon from the areas.

6.3 SOME FACTS ON BREAST CANCER

Micro-calcifications are tiny deposit of calcium in the breast, which causes cancer. It cannot be felt but can be detected on a mammogram. A cluster of these very small specks of calcium may indicate that cancer is present. Micro- calcification in the breast shows up in the breast as white speckles on breast X-rays. The calcifications are small, usually varying from 100 micrometer to 300 micrometer, but really may be as large as 2 mm. Though it is very difficult to detect the calcifications as such, when more than 10 calcifications are clustered together, it becomes possible to diagnose malignant disease.

Calcifications in general can be of two types--benign and malignant. Benign micro-calcification occurs when calcium is laid down within normal tissue in the breast. They do not spread to tissues around them or to other parts of the body. The presence of cancer is generally identified from the shape and pattern of the calcium specks. Malignant tumors can invade and destroy nearby tissues and spread to other parts of the body. Mostly, calcifications are composed of calcium phosphates in the form of hydroxyl apatite: less common are calcifications composed of calcium oxalate dehydrate. Small calcifications (< 12 mm) often referred to as micro calcification, are seen in about one third of all breast cancers and sometimes constitute the only radiographic sign of malignancy.

6.3.1 Identification of Micro-calcifications

The mammograms are all media-lateral oblique view, and are digitized with a scanning microdensitometer at a resolution of 50 micrometer × 50 micrometer. The region of interest in the original image is decomposed into a set of sub-bands of different frequency using multiresolution decomposition algorithm. The detection of micro-calcifications is accomplished by setting the wavelet coefficients of the low frequency sub-band to zero in order to suppress the image background information before the reconstruction of the image. After removing the background noise, each image is reconstructed by wavelet synthesisfiltering and up-sampling along the rows and columns of the image.

The bright point, which is present in the mammogram image taken, is separated out from its background. After the white top hat operation, pixels above the threshold are grouped together to form separate objects including Micro- calcifications as well as single pixels, noise and artifacts. Hence, only the objects above 9 pixels in the area are considered to be potential Micro-calcifications. This thresholding technique is followed by edge detection and corresponding region labeling and removal. Detection and diagnosis of breast cancer is a complex clinical problem and often results in false negative mammography readings. Breast cancer is among the most common and deadly of all cancers. Mammography is a uniquely important type of medical imaging used to screen for breast cancer. All women at risk go through mammography screening procedures for early detection and diagnosis of tumor.

A typical mammogram is an intensity X-ray image with gray levels showing levels of contrast inside the breast that which characterize normal tissue and different calcifications and masses. The contrast level of a typical mammogram image is proportional to the difference in X-ray attenuation between different tissues. Important visual clues of breast cancer include preliminary signs of masses and calcification clusters. A mass is a localized collection of tissue seen in two different projections, and calcifications are small calcium deposits. Unusually smaller and clustered calcifications are associated with malignancy while there are other calcifications that are typically benign. Such calcifications are termed as Micro-calcifications. In the early stages of breast cancer, these signs are subtle and hence make diagnosis by visual inspection difficult.

6.3.2 Background

The role image processing play in mammogram analysis is threefold: detection, diagnosis and noise cancellation. Detection involves identifying cancerous tissues in a mammogram. It provides spatial information about the Micro-calcifications. Some of the more important pitfalls in the diagnosis of breast cancer encountered are low contrast and poor image quality in mammography.

6.3.3 Diagnosis Tools

The diagnosis task is modeled as a two-class classification task. Features are extracted from Regions of Interest (ROIs - the region containing the masses or them Micro-calcifications) containing the abnormality, whose spatial information is provided by detection algorithms and each ROI is classified using a classification algorithm which is a supervised method that is first trained on a set of sample images (whose classification is known) called the training set. The performance of the algorithm is then tested on a separate testing set. Classifying a mammogram with Micro-calcifications is more challenging than with masses because of their erratic shapes, size, density and texture.

6.3.4 A Database of Mammograms and Preprocessing

Breast cancer is the most prevalent cancer and a leading cause of death in women today and about 6% of women develop the disease during their lifetime. As the cause of the disease is not clearly understood, primary prevention is not possible. However, since the current methods of treatment are quite effective against breast cancer in its early stages, early detection through mammograms, is the best way to reduce mortality from breast cancer .Micro-calcifications are deposits of calcium that are seen on mammograms or histological examination and are the earliest sign of the disease. Micro-calcifications appear as small white spots similar to grains of sand with a diameter of less on 0.5 mm and are grouped closely together to form clusters. They are often extremely difficult to detect since: (i) they are very small in size, (ii) inhomogeneous background, (iii) small calcifications or calcifications located in a dense breast parenchyma typically have a low contrast with the background, (iv)artifacts most often mimic micro-calcifications.

The MIAS database of digital mammograms contains left and right breast images of 161 patients. Its quantity consists of 322 images, which belongs to three types such as Normal, benign and malignant. The database has been reduced to 200 micron pixel edge, so that all images are 1024 x1024. There are 208 normal, 63 benign and 51 malignant (abnormal) images. It also includes radiologists ‘truth’ marking on the locations of any abnormalities that may be present. The database is concluding of four different kinds of abnormalities namely: architectural distortions, stellate lesions, Circumscribed masses and calcifications.

6.4 TEXTURE ANALYSIS METHODS FOR MAMMOGRAM

The early detection of breast cancer is done by a special type of x-ray called mammography. The classification of calcium deposits into the affected and non affected categories is a difficult task, which includes the detection of tumors as suspicious regions with a weak contrast to their background and the extraction of features which characterize malignant tumors. Moreover the detection of masses in mammograms are too sensitive and so it is better to use the supervised and unsupervised classification methods in mammography images in order to describe of various discriminate characteristics of both true and false phenomenon from the areas.

So far many researchers have gone for the mammography image classification to detect the calcium deposit in digital mammography classification and an unsupervised scheme for detection of micro-calcification on mammograms with the help of K-Means clustering algorithm seemed to be the difficult techniques because of their calculation that involved a lot of complex mathematical expressions. So it is inevitable to develop an unsupervised scheme for detecting micro-calcifications that uses features extracted using Gabor filtered images. Gabor filters are very effective in detecting fine textural patterns in mammograms and the unsupervised scheme for detection of micro-calcifications that does not make use of any a priori information as opposed to most of the techniques developed in the past. The aim in this section is to perform texture analysis on mammogram in both supervised and unsupervised methods.

6.5 PROPOSED TEXTURE SEGMENTATION FOR MAMMOGRAM

In the view of detecting micro-calcification on mammography images, the proposed work has two parts namely segmentation and classification. The first part starts with the selection of non-preprocessed of size in three levels 128×128, 256×256, and 512 ×512, which are then to be extracted in the form of detecting fine textural patterns. For the purpose of feature extraction, here Gabor filters have been chosen. Since the mammographic images frequencies are lying in the band pass frequency levels that match with Gabor filter frequency. Once the images are filtered and then to be subjected in to a histogram based thresholding to obtain binary images. The threshold (T) value can be found out by calculating the mean value which is to be added with the variance. Now the binary image is applied in the different texture analysis classification methods. The same set of procedure is to be applied in extracting the features from the mammographic images in the unsupervised classification method using the K-means clustering algorithm. The steps for the segmentation method are shown in Figure 6.1.

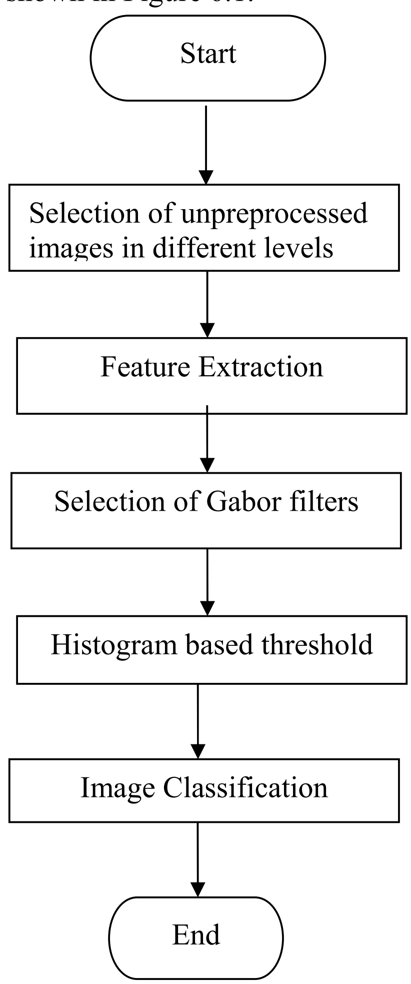


Figure 6.1 Image Classification Procedures

The analysis of the Mammogram is the need to identify the specific types of regions of interest. One way to achieve this is to develop features of the image which can be used to classify the image data. The greatest difficulty lies in finding some property of the image from which such features may be extracted. In discussing this it is helpful to consider the way in which pictorial scenes are interpreted. It is generally believed that one of the main visual cues is structured and difference in textural properties between regions. The texture of a region describes the pattern of spatial variations of gray tone in a neighborhood, where the neighborhood is small compared to the region particularly for images of non homogeneous textures such as mammograms and an additional step of segmentation or classification is required.

6.5.1 Region of Interest Identification (ROI)

The first stage of micro-calcification detection is Region of Interest (ROI) Identification. The enhanced mammogram image is decomposed by filters. The resulting horizontal detailed image and vertical detailed image is used to identify the region encircling the micro-calcification clusters. Third and fourth order statistical parameters Skewness and Kurtosis are used to find the regions of microcalcification clusters. An estimate of the Skewness is

$$S_k = \frac{\sum_{i=1}^{N}(x_i - \overline{m})}{(N-1)\sigma^3} \qquad (6.1)$$

And the statistical parameter Kurtosis holds the expression

$$K_u = \frac{\sum_{i=1}^{N}(x_i - \overline{m})^4}{(N-1)\sigma^4} \qquad (6.2)$$

where x_i is the input data over N observations, $\overline{m}$ is the ensemble average of x_i,' $\sum$ ' is the standard deviation.

The third and fourth order statistical estimates are calculated for every overlapping 32×32 square regions of horizontal band pass image or vertical band pass image. The area having skewness value greater than 0.2 and kurtosis value greater than 4 is marked as region of interest. Micro-calcification appears in mammogram as modular points with higher brightness, localized are broadly diffused along the breast tissue; where as normal tissues such as blood vessels are linear in structure. So detecting the modular structure in image is a key in detecting the micro-calcification.

6.5.2 Nodular Extraction

After identifying the ROI, the next step is to detect edges of the enhanced image. Edges are detected by applying the first order derivatives. First order derivatives are implemented by gradient operator Gx and Gy. The edges are detected by computing the gradient of each pixels in the enhanced image in X and Y direction. By applying the first order derivatives enhanced edges are detected. These enhanced edge pixels are connected in sets to form group. The group containing less than 5 pixels discarded. Final goal is to extract only circular (nodular) region whose radius is equal to micro-calcification radius. After obtaining different circular region micro-calcification detection is performed to distinguish between valued micro-calcification regions and invalid one. The two validity measurement parameters are cluster density and relative shell thickness. The two parameters are used to identify micro calcification regions.

$$C_d = \frac{\sum_{k=1}^{n} U_{tk}}{2\pi r_t} \qquad (6.3)$$

Numerator denotes sum of membership function on most characteristics points $|U_{tk}|$ >1/2 and denominator denotes area of the circle region. Relative shell thickness is defined by

$$RSTt = \frac{\sum_{k=1}^{n} U_{tk}^{m} D_{tk}^{2}}{r_t \sum_{k=1}^{n} U_{tk}^{m}} \quad (6.4)$$

where r_i is radius of circular region. Correct micro-calcification is detected according to the rule If (C_{di}>1.15 and RST_i <0.2).

6.5.3 Gray Level Co-Occurrence Matrix

Grey-Level Co-Occurrence Matrix (GLCM) as one of the most widely used statistical texture measures. The idea of the method is to consider the relative frequencies for which two neighboring pixels are separated by a distance on the image. Since the GLCM collects information about pixel pairs instead of single pixels, it is called a second-order statistic. Texture measures, such as homogeneity, contrast, and entropy are derived from the co-occurrence matrix. The different sets of images of MIAS have been tested and compared.

6.5.4 Texture Spectrum Operator

The statistics of the frequency of occurrence of an image should reveal texture information and texture spectrum is sensitive to the directional aspect of texture. The undesirable influence of the regional intensity background is eliminated from the texture spectrum. Here sample images from Mammographic Image Analysis Society (MIAS) data base have been taken and the optimal window size is selected for the further classification.

6.5.5 Local Binary Pattern Operator

Local Binary Pattern Operator does not take into account the contrast of texture which is the measure of local variations present in an image and is important in the description of some textures. Texture spectrum operator is similar to LBP Operator but it uses three levels that is, two thresholds instead of two levels used in LBP Operator. This leads to a more efficient representation and implementation than with LBP Operator and according to experimental tests with the help of varying the widow size for different same set of mammographic images that were used in TSO and EBLD methods.

6.5.6 Entropy Based Local Descriptor

The main objective of Entropy based local descriptor method is for the texture measure widely used to quantify the smoothness of image texture since Entropy does not depend on actual values in texture. High entropy based local descriptor is associated with a high variance in the pixel values, while low entropy based local descriptor indicates that the pixel values are fairly uniform. Here 60 sample images have been tested to detect the suspicious masses on mammographic images.

6.6 SIMULATION AND RESULTS

6.6.1 Simulation Study

Here Texture Spectrum Operator (TSO), Entropy Based Local Descriptor (EBLD), Local Binary Pattern Operator (LBP) and Gray Level Co-occurrence Matrix (GLCM) have been evaluated for the same set of MIAS images. The results for classification accuracy have been computed and compared with the different supervised and unsupervised methods. The classification accuracy for Texture Spectrum operator (TSO) and Entropy Based Local Descriptor (EBLD) operator achieve less accuracy as compared with Local Binary Pattern operator (LBP) and Gray Level Co occurrence Matrix(GLCM). The images have been taken from the Mammogram Image Analysis Society (MIAS) for the validation of the proposed algorithm. The images used are having 512×512 pixels. Different sizes of range blocks chosen are 16×16, 32×32 and 64×64 pixels and the optimal value fixed is 16 to maximize the classification accuracy. The performance of the proposed segmentation algorithm has been evaluated by calculating three parameters namely True Positive (TP), False Positive (FP) and False

Negative (FN) and it is found that the proposed algorithm is so attractive in terms of TP, FP and FN compromising the classification accuracy.

6.6.2 Classification

The classification of masses from normal regions requires high quality classification system as most of these differences can be at times subtle. In order to determine the discriminating effectiveness of texture features extracted from co-occurrence matrices constructed at different distances (each distance yields a different feature set), classification is initially performed using each normalized feature set and linear discriminant analysis. In the proposed investigation, 60 breast classification images have been used and tested with different supervised method to evaluate classification accuracy.

In the analysis of results within the study, the following definitions are used:

- True Positive (TP): lesions called cancer and prove to be cancer.
- False Positive (FF'): lesions called cancer that proves to be benign.
- False Negative (FN): lesions that are called negative or benign and prove to be cancer.
- True Negative (TN): lesions that are called negative and prove to be negative.

To evaluate the performance of the classification methods True Positive (TP) and False Positive (FP) can be defined as:

FP=FP+TN

TP =- TP

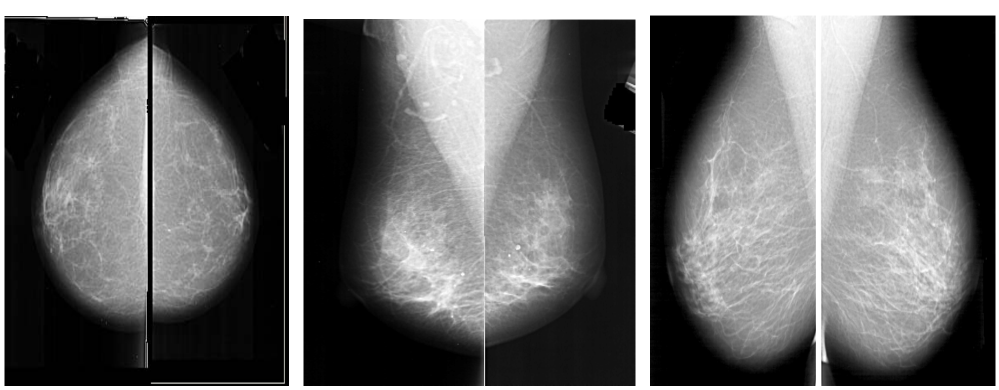
FP=-TP+ FN

6.6 3 Location of a Common Reference Point

In order to align left and right breast image pairs prior to bilateral subtraction a common reference, the spatial position of the nipple, is located.

6.6.4 Alignment and Bilateral Subtraction of Left and Right Breast Images

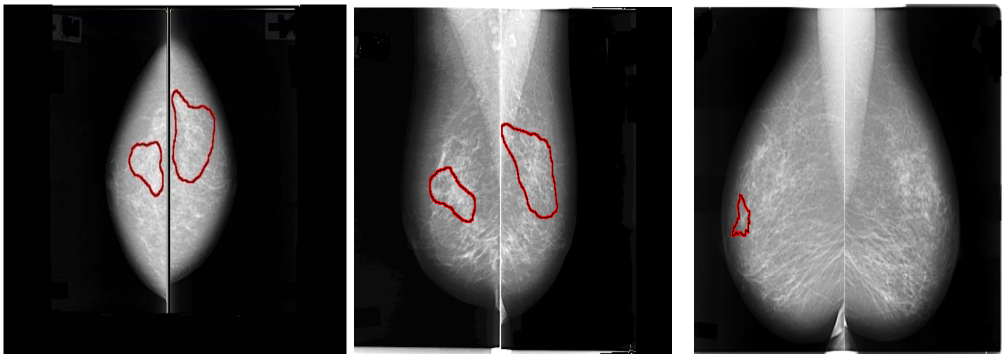
Once the nipple has been located in each breast image of a matching pair, the observed image is translated such that the nipple locations of both breast images are aligned. Taking different rotations of the observed image and determining a correlation measure against the reference image, the left original breast image, the best alignment can be determined. Initially the observed image is subjected to large degrees of rotation, following which the best range is located and the process repeated with smaller incremental rotations until the best correlation is found. Using aligned left and right breast image pairs, two images are generated by bilaterally subtracting one image from the other. One is a positive image and details differences that occur in the left breast image and not in the right, and a negative image that details differences that occur in the right breast image and not in the left. The sample suspicious image as shown in Figure 6.2 (a) to (c) have been taken from the MIAS data base for the purpose of image classification.The sample image were allowed to the GLCM,LBP,EBLD and TSO for the classification.



(a) (b) (c)

Figure 6.2 (a) to c) Original Mammogram Suspicious Images

Figure 6.3 (a) to (c) shows the output image after applying GLCM classification method, where the micro-calcification could be easily found and the same image was tested with the TSO and EBLD methods.



(a) (b) (c)

Figure 6.3 (a) to (c) Detected Micro-calcifications After Texture Analysis using GLCM

Figure 6.4 shows the comparative analysis between TSO and EBLD and it is studied that the EBLD exhibits better classification rates, in terms of identifying True Positive (TP).

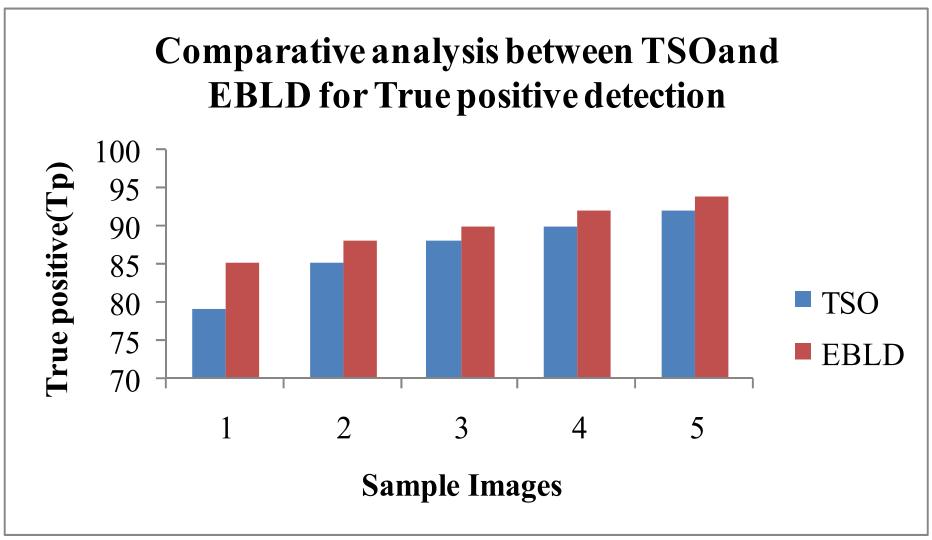


Figure 6.4 Comparative Analyses Between TSO and EBLD for True Positive Detection

Figures 6.2 and 6.3 show the normal breast cancer images from the MIAS data base and those have been tested with the classification methods followed by the proposed segmentation method. From Figure 6.4, it is identified that EBPLD method has high classification accuracy while finding the micro-calcification of mammogram in terms of true positive compared with TSO method.

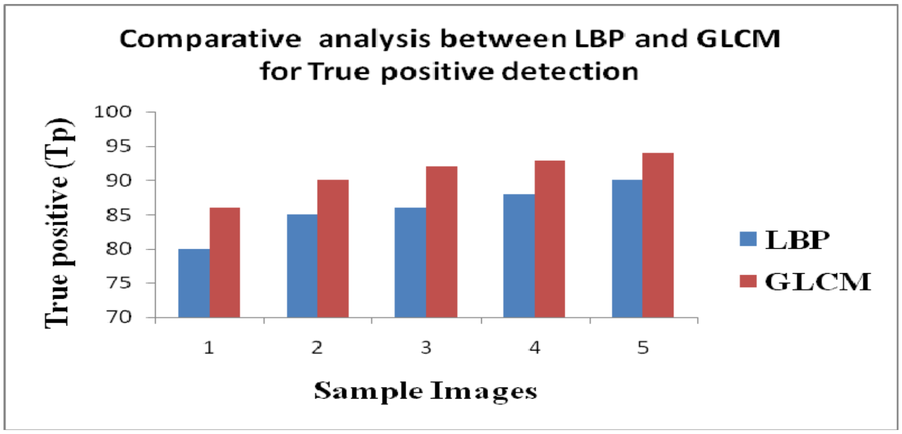


Figure 6.5 Comparative Analyses Between LBP and GLCM True Positive Detection

From Figure 6.5, it is found to know that GLCM method is so attractive than the LBP method in true positive.

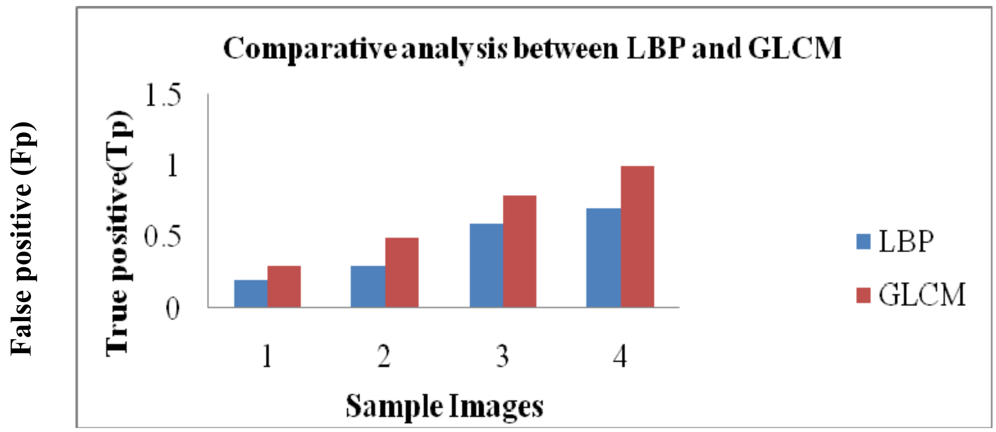


Figure 6.6 Comparative Analyses Between LBP and GLCM False Positive Detection

Likewise Figures 6.6 and 6.7 shows that the comparative analysis of four supervised methods in terms of finding false positive mammogram and from the experimental results, it is found that the GLCM method outperforms the other methods.

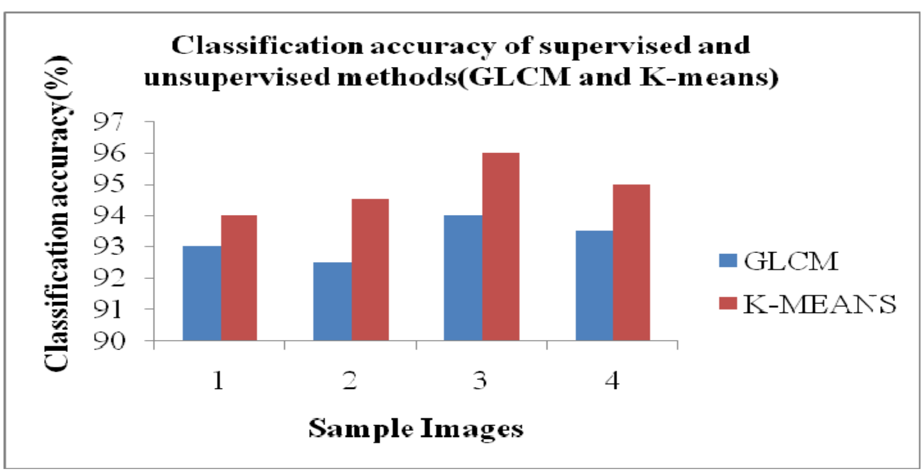


Figure 6.7 Comparative Analysis Between GLCM and K-Means

The over all results show that the sensitivity has been 95% in classification accuracy through supervised method in GLCM followed by LBP, EBLD and TSO.

6.6.5 Algorithms Used for Comparison

In the comparative study, true positive rate is the ratio for the correctly detected micro-calcification clusters and false positive is the number of detection for the breast detection, which is shown in Table 6.1. The results show that the GLCM method takes an edge in identifying the true positive and false positive. Table 6.2 shows the comparative analysis of different supervised classification methods.

Table 6.1 Comparison Table for True Positive and False Positive

Different classification Methods	True Positive TP (%)	False Positive FP(%)
TSO	82.00	0.56
EBLD	90.00	0.78
LBP	92.00	1.04
GLCM	93.5	2.15

After finding the classification rates by applying different texture analysis method, the results have been compared.

Table 6.2 Comparative Analysis of Different Classification Methods of Supervised and Unsupervised Methods

<table>
<tr><th colspan="2">Different classification methods</th><th>Percentage
of mass detection
(Classification accuracy)%</th></tr>
<tr><td rowspan="4">Supervised classification methods</td><td>TSO</td><td>82.00</td></tr>
<tr><td>EBLD</td><td>90.00</td></tr>
<tr><td>LBP</td><td>92.00</td></tr>
<tr><td>GLCM</td><td>93.5</td></tr>
<tr><td>Unsupervised classification methods</td><td>K-Means</td><td>95.00</td></tr>
</table>

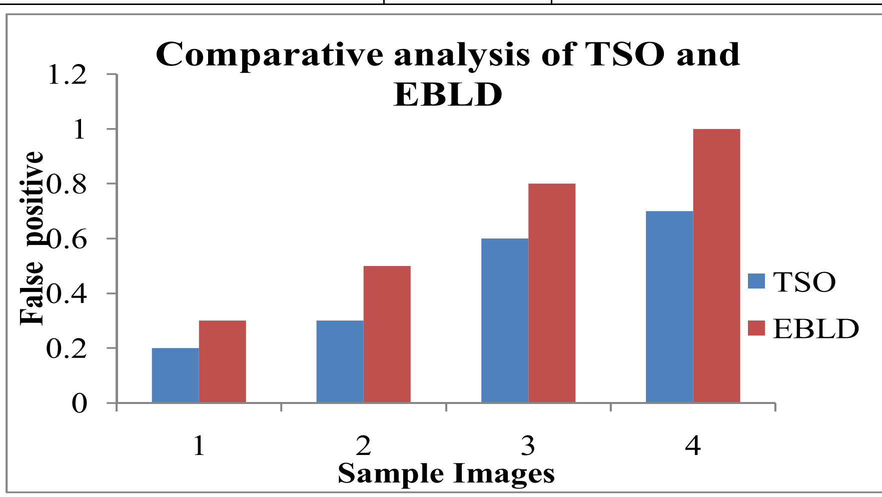


Figure 6.8 Comparative Analysis of Number of Iterations Versus Fp for TSO and EBLD

From the Figures 6.8 and 6.9, it is understood that EBLD and GLCM have the edges in terms of detection of false positive.

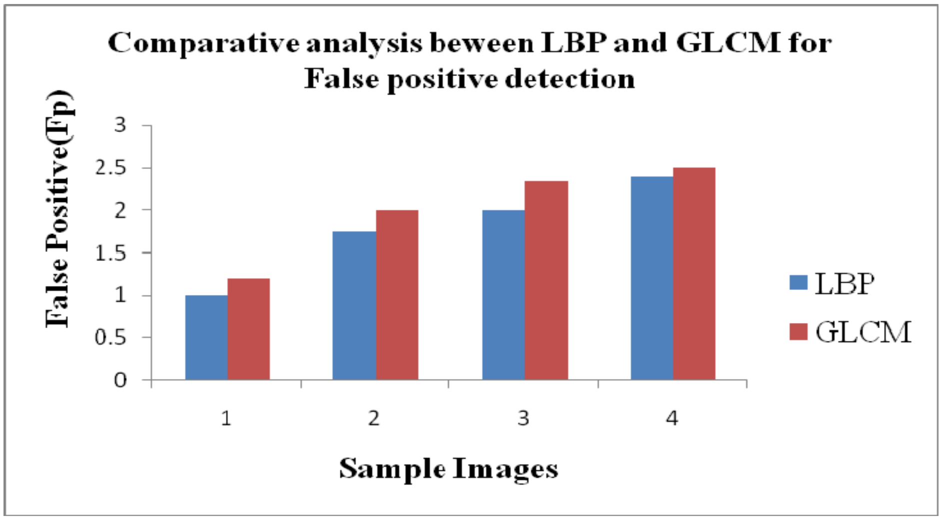


Figure 6.9 Comparative Analysis of Number of Iterations Versus Fp for LBP and GLCM

From the experimental results, it is clearly found that the classification accuracy of GLCM is overtaking other supervised methods. It is also clearly found that the classification accuracy of K-means method has higher classification accuracy than the GLCM and it is concluded that the unsupervised method is more suitable than the supervised method for the purpose of finding micro-calcification on mammograms. The images have been taken from the mammogram image analysis society (MAIS) for the validation of the proposed algorithm to compare different supervised classification methods. The images used are having 512×512 pixels. Different sizes of range blocks have been chosen are 16×16, 32 ×32 and 64×64 pixels and the optimal value fixed is 16 to maximize the classification accuracy. The performances of different classification methods have been evaluated by calculating three parameters namely TP, FP and FN and it is found among the different supervised classification method GLCM is so attractive than the other methods to detect micro- classification for mammographic images.

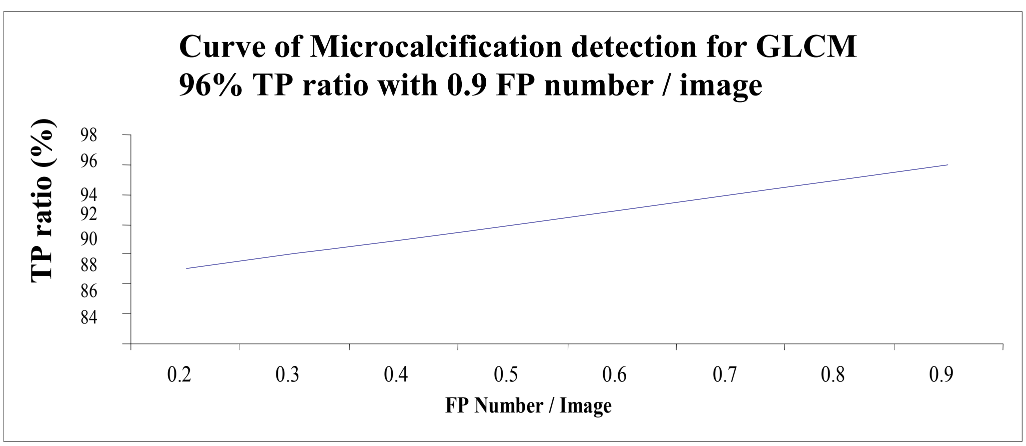


Figure 6.10 Curve of Micro-calcification detection for GLCM

The Figure 6.10 represents the curve of Micro-calcification detection for GLCM of 96% TP ratio with 0.9 FP number per image.

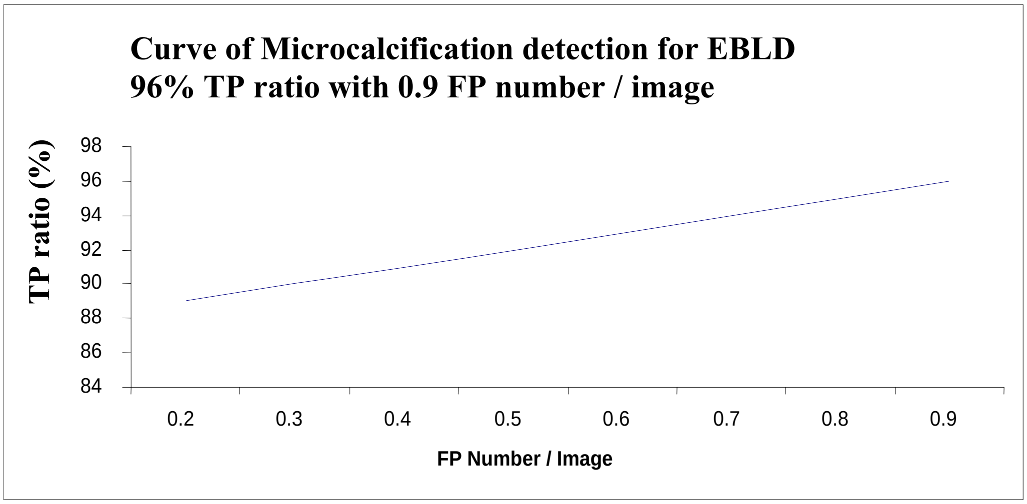


Figure 6.11 Curve of Micro-calcification detection for EBLD

The Figure 6.11 represents the curve of Micro-calcification detection for EBLD of 94% TP ratio with 0.9 FP number per image.

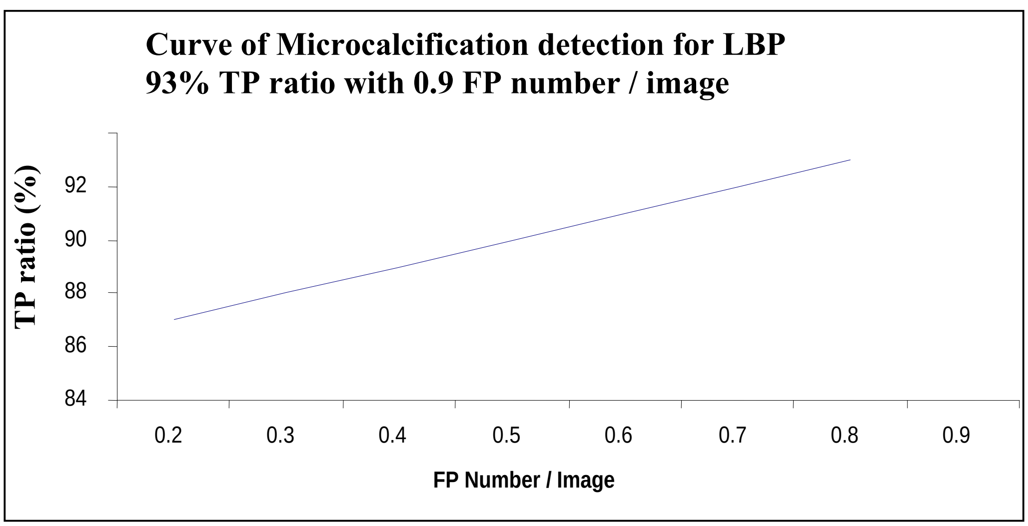


Figure 6.12 Curve of Micro-calcification detection for LBP

The Figure 6.12 represents the curve of Micro-calcification detection for LBP of 93% TP ratio with 0.9 FP number per image.

6.7 DISCUSSION

This section presents the performance results of the proposed supervised and unsupervised methods to detect the calcium deposit and compared with various existing methods such as obtained through simulation. The results are measured in terms of true positive, false positive and false negative while calculating the micro-calcification.

In the investigation 60 breast classification images have been used and tested with both supervised and unsupervised methods to evaluate classification accuracy. The results show that the sensitivity has been 93.5% in classification accuracy through supervised method and 95% through unsupervised method. The Figure 5 shows that the original mammography images and the Figure 6 shows the detected micro-calcifications from the mammogram images for the supervised image classification.

6.8 SUMMARY OF CONTRIBUTIONS

This chapter presents supervised and unsupervised methods to find out the classification accuracy of mammogram, which focuses on different aspects, namely, true positive, fast positive, fast negative in the view of confirming the breast cancer. Since breast cancer is one of the major causes for the increase in mortality among middle aged women, especially in developed countries. The X-ray mammography is the most common technique used by radiologists in the screening and diagnosis of breast cancer in women. Mammography associated with clinical breast examination is the most efficient method for early detection of cancer. It is the best examination technique for the early defection of breast cancer reducing the mortality rates by up to 25%, their interpretation requires skill and experience by a trained radiologist.

The aim of this study is to analyze digitized mammograms by applying computer Image processing techniques to enhance X-ray images and subsequently extract features from suspicious regions characterizing the underlying texture of breast region. The presence of micro-calcifications (tiny granule like deposits of calcium) in X-ray mammograms is considered an important indicator for the detection of breast cancer. Here in this chapter the new improved segmentation algorithm for classification has been introduced. The proposed algorithm has been applied in different classification methods and it is concluded that K-Means unsupervised classification method performs well in terms of finding out micro-calcification than the other methods.

In the next chapter, the conclusion, summary of discussion and the future scope are explained.

CHAPTER 7
CONCLUSIONS AND FURTHER SCOPE

7.1 CONCLUSIONS

In this monograph, the uses of different statistical methods to improve the image output by means of suitable methods and detection of mammogram have been concluded. UHCTA, PBCM and new segmentation algorithms for supervised and unsupervised have been applied in various images has been discussed.

Much work has been proposed in literature over the past few decades. Works related to classification problem have been surveyed and are presented in the monograph. The literature survey is presented to illustrate that more work on similar problems have been carried out. Various literatures on improving classification accuracy of images by both supervised and unsupervised methods have been briefly stated in Chapter **2** of this monograph.

The features of textures using supervised classification methods were studied. The local texture information for a given pixel and the neighborhood is characterized by the corresponding texture unit. The global textural aspects of an image are revealed by its texture spectrum. Here supervised classification method has been used to extract the textural information of an image with respect to texture characteristics. The purpose of this phase is to present a new statistical method of texture analysis which is focused on texture characterization and discrimination with features like local binary pattern operator, texture spectrum operator and entropy based local descriptor. The obtained results are provided in Chapter **3** of the monograph.

The postulation of the proposed UHCTA algorithm is compared with local binary pattern operator and it is found that the proposed method increases in classification accuracy. The proposed UHCTA algorithm is also compared with local binary pattern operator, gray level co-occurrence matrix; entropy based local descriptor and K-Means. From the evaluation results it is found that the UHCTA method out performs the other classification methods. According to the proposal, it is also observed that the accuracy can be increased while increasing the window size. The obtained outputs of the proposed methods are provided in Chapter **4** of the monograph.

When using statistical approach in texture analysis for image classification more complication is to be met. Particularly grey level co-occurrence matrix approach is applied in discriminating different texture images that compromises the classification rate. So it is necessary to concentrate on image classification method based on pixel by pixel with maximum likelihood estimates that must be compared to a single window classification not only to monochrome images but with the color images too. The proposed method shows with attractive results in terms of classification accuracy and some of the tested images are also provided in Chapter **5**.

Texture analysis has been very much used in medical image problems. Among all medical image tasks detecting the micro-calcification on mammogram is the most difficult one. Since breast cancer is the most prevalent cancer that leads to death in woman today. Moreover micro-calcification are deposits of calcium that can be seen in mammogram and owing to the small size of micro-calcification with a diameter of less than 0.5mm level and are in the form of groups as clusters in homogeneous background it is very difficult to detect. In Chapter 6, using supervised and unsupervised methods finding the suspicious masses on mammogram were studied and the comparative analysis shows that the GLCM works better than the K-Means.

The results for detection of mammogram by means of classifying the suspicious images with GLCM and K-means are analyzed and the comparative study has been done by taking various images. Analysis of the outputs of supervised and unsupervised approach shows that the proposed algorithm provides better results. It is concluded that the classification with GLCM performs better than the K-means. In addition, with the proposed UHCTA provide better results when compared with existing approach for improving classification. The results of proposed classifications are encouraging. Some of the comparative tables and graphs are provided. In overall comparison, the proposed algorithms of both supervised and unsupervised methods provide better results when compared with existing techniques on improving classification accuracy.

The following conclusions were attained from the present study and are as follows.

- Supervised classifications of textural features have been studied to understand the concept of texture analysis.
- Improving classification accuracy using an unsupervised hybrid method has been studied by integrating different combination of features from different classification methods.
- The classification accuracy could be well possible by using individual features rather than combination of features.
- Succeeded in maximizing the classification accuracy by using variable window sizes instead of fixed window size.
- Classification on color images of statistical texture images have been studied to reduce the difficulty in classifying the typical color images.
- Using the method of pixel by pixel, the color images can easily be classified with the concept of overlapping of neighbor pixels.
- Comparative analysis of different statistical texture analysis has been studied to identifying the advantages and disadvantages of each method.
- Applying the statistical methods in medical images particularly on mammograms to determine the calcium deposit effectively.

The overall results suggest that the studies on statistical texture analysis and its applications impart a better idea in the way of attaining remarkable classification accuracy.

7.2 SUGGESTIONS FROM THE PRESENT STUDY

The concepts are discussed in detail in various sections of this monograph. There are various application areas for texture analysis. One of the important application areas of this texture analysis has been identified as mammogram in medical application, and many problems related with routing have been explored in this monograph. The conventional approaches for improving classification accuracy for texture analysis may not be good in many ways. Here, a proposal has been given to show that the problems in conventional approaches avoided, if the proposed classification methods with different modules and a novel classification approach is adapted for improving image for many applications.

The results suggest that the proposed classification methods for texture analysis perform significantly better in the way of incorporating remarkable accuracy measures with reliability. The approach would definitely increase its usability and adds more to the future prospects of improving classification in texture analysis for many applications.

7.3 OPEN ISSUES AND FUTURE WORK

In this study, the different methods of statistical texture analysis have been analyzed and the comparative analysis also has been illustrated. Even though the results are so attractive in terms of reliability in classification accuracy, the supervised and unsupervised methods have to be tested with features in order to improve the texture analysis quality for the image classification. Moreover the color image features like homogeneity, contrast are also to be taken in to consideration for the texture classification since color images classification seems to be complex process. Texture analysis is moving towards further technical advancements in both logically and accurately in the digital image processing. So there is an obvious need to strengthen the existing ones. Even though the classification methods of different modules designed in this monograph is extremely good for texture analysis, it can be upgraded in future upon considering different features. There are a number of increased applications that require many new parameters such as contrast, autocorrelation, angular moments, inertia, means, and standard deviation and so on mainly to increase the classification accuracy.

REFERENCES

[1]Fundamentals of Global Positioning System Receivers: A Software Approach by *James Bao-Yen Tsui.*

[2]Chheda Gaurav, Gajra Niket, Chhaya Manal, Deshpande Jitesh, Gharge Saylee, "Real Time Bus monitoring and Passenger Information System", vol. 1, no. 6, pp. 2231-2307, January 2012.

[3]R. Maruthi, C. Jayakumari, "SMS based Bus Tracking System using Open Source Technologies", *International Journal of Computer Applications (0975–8887),* vol. 86, no. 9, January 2014.

[4] Swati Chandurkar, Sneha Mugade, Sanjana Sinha, Megharani Misal, Pooja Borekar, "Implementation of Real Time Bus Monitoring and Passenger Information System", Volume 3, Issue 5, May 2013.