Classification of Mammograms based on Gray Level Intensity

A Dissertation

Submitted in partial fulfillment of the requirements for the award of degree of

Master of Engineering

In

Electronic Instrumentation and Control

Submitted By:

Pratham Arora

(801151019)

Under the esteemed guidance of

Mr. Mandeep Singh
Assistant Professor
Department of Electrical and Instrumentation Engineering

ELECTRICAL AND INSTRUMENTATION ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

(Established under the section 3 of UGC act, 1956)

PATIALA, 147004, Punjab, India

JULY 2013

DECLARATION

I hereby certify that the work which is being presented in the dissertation entitled as **"Classification of Mammograms based on Gray Level Intensity"** in partial fulfillment of award of degree of **Master in Engineering in Electronic Instrumentation and Control** submitted in Electrical Instrumentation Department, Thapar University, Patiala, is an authentic record of my own work carried out under the supervision of **Mr. Mandeep Singh**, Assistant Professor, Department of Electrical Instrumentation Engineering, Thapar University, Patiala.

Date 15/7/13

Pratham Arora

(801151019)

I certify that the above statement made by the student is correct to the best of my knowledge and belief.

15/7/13

Date

(Mr. Mandeep Singh)
Assistant Professor, EIED
Thapar University,
Patiala

Countersigned by

Dr. Smarajit Ghosh
Head of Department, EIED
Thapar University
Patiala, Punjab

Dr. S.K. Mohapatra
Dean of Academic Affairs
Thapar University
Patiala, Punjab

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my supervisor, **Mr. Mandeep Singh,** Assistant Professor (EIED), for all his guidance and invaluable advice throughout the progress. He has stimulated my interest in medical image processing and inspired me for doing research on this topic.

I would also like to thank **Dr. Smarajit Ghosh**, Professor & Head, Electrical and Instrumentation Engg. Department, and **Dr. S.K Mohapatra**, Dean, Academic Affairs for giving an opportunity to work in this regard.

Also I would like to thank my parents and my friends who were always there for me. It would not have been possible without their continuous support and encouragement.

(Pratham Arora)

ABSTRACT

Breast cancer is one of the leading causes of death among women. To find out the type of breast tissue is the first step towards knowing whether a woman has cancer /likely to develop the risk of cancer or not. Out of fatty, glandular and dense breast tissue types; dense tissue type has the most probability of having/developing malignant tissues. In this work, mammograms are treated as textures and a method is proposed to classify them into fatty, glandular and dense breast tissue type. Three new features based on gray level intensity are proposed in this work and are merged with well-known gray-level co-occurrence matrix (GLCM) features for tissue classification. In total 16 features are extracted from 150x150 region of interest (ROI). Out of 16 features, optimal features are selected through correlation feature selection (CFS). Finally, classification is done using Naïve Bayes classifier, Random Forest classifier and k-nearest neighbor (k-NN) classifier on four sets of features which are: only GLCM features, three proposed features, combination of both (GLCM+proposed) and seven optimal features. The results of all the set of features are compared on the basis of their accuracy. Subsequently, mammograms are classified into fatty and dense using the same procedure.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 OVERVIEW

Cancer has been like a curse to the human race for years and over next few decades it is likely to become the leading cause of death. According to the stats from World health organization (WHO), cancer lead to the 13% of all the deaths of the world in 2004 and this number is estimated to grow in the future [1]. Breast cancer is a major concern of health as it is the second leading cause of death in the western countries among women [2]. A report from the National Cancer Institute (NCI) states that in the United States from every 8 women 1 (approximately 12.6%) will develop breast cancer in their lifetime [3]. Presently, there are no efficient means of preventing breast cancer because its cause is still a mystery. But effective diagnosis and treatment at early stages can make women recover from the cancer.

Breast with higher density generally points towards the presence of malignant tissue. Sometimes it becomes difficult to detect malignant tissue because the higher breast density tends to mask the abnormal tissues. Due to this masking effect higher breast density can cause problems while detecting malignant tissues [4]. It has been shown in [5] that there is a strong correlation between mammographic breast density and risk of developing cancer. Higher density breasts are more liable towards the risk of developing breast cancer. It should be noted that, while in this dissertation, we refer to breast tissue density and dense tissue, in the wider literature the following terms are also in use: parenchymal patterns, fibro glandular disk, and parenchymal density.

Texture is a significant feature of many image types, which is the form of information/data or arrangement of the structure found in a picture. Texture feature is a sort of visual characteristics that does not depend on color or intensity and speculates the intrinsic phenomenon of images. It is the total of all the intrinsic surface properties. That is why the texture features have been widely used in image processing. Many objects in an image can be distinguished purely by their textures without any other information [6].

1.2 OBJECTIVE OF THE DISSERTATION

The objective of this dissertation is:

- Classification of mammograms into fatty, glandular and dense tissue type.
- Check out the accuracy of the features proposed in this dissertation with respect to the other features for tissue classification.
- Find out the best classifier out of Naïve Bayes classifier, Random Forest classifier and k-nearest neighbor (k-NN) classifier for classification.

1.3 ORGANIZATION OF THE DISSERTATION

The structure of the dissertation is as follows:

Chapter 1 deals with some of the facts about cancer and breast cancer; brief introduction about breast tissue density type and a brief introduction about the objective of the work.

Chapter 2 involves the literature survey, which explains the various works that have been carried out in this field.

Chapter 3 gives a brief introduction of mammograms and digital mammography.

Chapter 4 presents the statistical approach towards the texture analysis; use of GLCM features in texture analysis and the extraction of GLCM features from a particular texture.

Chapter 5 shows the database used in this work; explains about the preprocessing of the data, classifiers used and the classification techniques used to complete the analysis.

Chapter 6 presents the experiments done and the results of the experiments.

Chapter 7 concludes the work with the conclusion and the future scope in this area.

Chapter 2

Literature Review

2.1 TEXTURE DESCRIPTORS APPLIED TO DIGITAL MAMMOGRAPHY

This chapter deals with the discussion of the work done by many researchers' in the field of Digital Mammography. It gives a study on their contributions and publications in the same field. The study on digital mammography and breast cancer is going on for the betterment of mankind, and the objective is to improve the diagnosis of breast cancer patients and help the radiologists' for the treatment of the patients. The increase in the work in this area and the publications on the various and diverse methods was hence of prime importance.

In the early 80's Juhl, premised some particular features assayed in the regular tests as usual indicants of malignancy and it reckoned a huge growth and increment of investigating the breast cancer. Small masses and micro calcifications are the widely known signs in this field of digital mammography [7]. In the late 80's, researchers like Giger by the ways of asymmetry studies worked on the methods for finding and examining mammographic masses which contributed a lot in this field [8].

In the early 90's researchers like Davies and Dance, Chan *et al.*, Karssemeijer *et al.* also contributed their part in the field of digital mammography and introduced their works on the detection of micro-calcifications [9][10][11]. It was the study done by Miller and Astley that set the benchmark in this field as it introduced the concept of texture analysis utilized to classify breast tissues. The immense growth of related works in this field introduced by new techniques and advancements in medical area was decisive thanks to the analysis of texture, which it is nowadays gaining importance worldwide [12].

In the late 90's Wei *et al.* investigated a linear discriminant classifier using the multi-resolution texture features to effectively classify masses from normal tissue on mammograms. With texture features based on the wavelet coefficients and variable distances, the average area, *Az*, under the ROC curve, reached 0.89 and 0.86 for the training and test groups, respectively. Wei *et al.* also investigated the application of multi-resolution global and local texture features to decrease false positive detection in a computerized mass detection program. The results of that investigation pointed out the effectiveness of the combined global and local features in the classification of masses and normal tissue for false-positive reduction. With both global and local features, the area, *Az*, under the test ROC curve, peaked to 0.92 for the manual dataset and 0.96 for the hybrid dataset, demonstrating statistically significant improvement over those obtained with global or local features alone [13] [14].

Ole *et al.* presented a new automated method for detection of tumors in digital mammograms based on the application of multichannel-filtering for texture feature extraction [15]. Measures of the skewness of the image brightness histogram, and measures of image texture characterized by the fractal dimension were investigated by Byng *et al.* to analyze film-screen mammograms. The correlation between both the measure was found to be very strong with radiologists' subjective classifications of mammographic parenchyma (Spearman correlation coefficients, Rs = -0.88 and -0.76 for skewness and fractal dimension measurements, respectively) [16].

In the twenty first century a discrimination of breast density implemented by Bovis and Singh was based on the underlying texture contained within the breast tissue apparent on a digital mammogram and realized by utilizing four approaches to quantify the texture. The testing data set was split into four categories: (a) predominantly fat; (b) fat with some fibro-glandular tissue; (c) heterogeneously dense; (d) extremely dense [17]. To discriminate lesions from normal tissues characteristics such as intensity, contrast, isodensity, location and texture were defined and tested by Brake *et al.* in [18].

Oliver *et al.* segmented the breast area into fatty versus dense mammographic tissue, extracted morphological and texture features from the segmented breast areas and then used a Bayesian combination of a number of classifiers. The evaluation showed a strong correlation (κ = 0.81 and

0.67 for the two different data sets) between automatic and expert-based Breast Imaging Report and Data System (BIRADS) mammographic density assessment [19].

Wavelet transform was investigated by Docusse *et al.* to classify micro-calcification borders and the results showed Symmlets wavelet family presented the best results with a 94% efficacy in their tests [20]. A curvelet transform based texture analysis was presented by Eltoukhy *et al.* for the classification of tissues [21].

Saidin *et al.* applied graph cut technologies to segment a mammogram into different mammographic densities and extended their work using seed based region growing techniques in to evaluate the graph cut techniques in the segmentation of the mammogram. In this dissertation the focus is on Grey-level Co-occurrence matrix (GLCM) [22] [23].

2.2 GRAY-LEVEL CO-OCCURRENCE MATRICES APPLIED TO MAMMOGRAPHY

Gray-Level Co-occurrence Matrix (GLCM) is one of the most widely used texture descriptors used in the field of digital mammography. Bovis and Singh in their work showed how textural features can be used in order to detect masses in mammograms employing five co-occurrence matrices statistics extracted from four spatial orientations, horizontal, left diagonal, vertical and right diagonal corresponding to (0^0, 45^0, 90^0 and 135^0) and four pixel distance (d=1, 3, 6, 9). Hence, a classification is performed using each feature vector and linear discriminant analysis [24]. Marti *et al.* in their work showed GLCMs are frequently used in computer vision obtaining satisfactory results as texture classifiers in different applications. So, it is useful in order to obtain more information about similarity between two images. Their approach uses mutual information with the purpose to calculate the amount of mutual information between images using histograms distributions obtained by grey-level co-occurrence matrices [25].

Blot and Zwiggelaar proposed two approaches based in detection and enhancement of structures in images using GLCM. The first one, the GLCM of the local region of interest is compared to a mean GLCM obtained from a number of equal size areas surrounding the local region of interest (ROI). The purpose is to compare the difference between these two matrices obtaining a probability estimate of the abnormal image structures in the ROI. In this sense, the second approach follows the same proposal extracting background texture in mammographic images with an improvement respect to previous works and it does not depend on any prior assumptions about the type of structures to be enhanced. Another study based on background texture extraction for classification of mammograms. Blot and Zwiggelaar presented their work where there is a statistical difference between GLCM for image regions that include image structures and regions that only contain background texture which is provided by a classification in mammograms. So, the classification of mammographic parenchymal patterns can be improved if anatomical structure can be removed from the image [26] [27] [28].

Some various approaches on the basis of co-occurrences matrix as a feature descriptors extraction were developed. A neuro-fuzzy model for fast detection of candidate circumscribed masses in mammograms and texture features are estimated using co-occurrence matrices which are used to train the neuro-fuzzy model was presented by Youssry *et al.* [29]. Whereas, Marti *et al*. proposed a supervised method for the segmentation of masses in mammographic images using texture features which present a homogeneous behavior inside the selected region [30].

Jirari proposes an intelligent Computer-Aided Detection system (CAD) constructing five co-occurrence matrices at different distances for each suspicious region. A different number of statistical features are used to train and test the Radial Basis neural network [31]. Another work is presented by Lena *et al.* with the study of the differentiation of dense regions containing speculated masses from regions of normal dense tissues, by means of feature analysis on wavelet-processed mammograms. A multi-resolution texture feature of second order statistics were extracted from spatial GLCM using different orientations and distances [32].

Lyra *et al.* study how to identify breast tissue quality data quantification using a CAD system, where images categorized using the BIRADS breast density index. The texture features were derived for each sub-region from an averaged gray- level co-occurrence matrix (GLCM) [33].

Karahaliou *et al.* investigate texture properties of the tissue surrounding micro-calcification (MC) clusters on mammograms for breast cancer diagnosis on Gray-level texture and wavelet coefficient texture features at three decomposition levels are extracted from surrounding tissue regions of interest [34] [35]. Panchal *et al.* used grey-level based image features and BIRADS lesion descriptors along with patient age and a subtlety value (radiologists' interpretation) for the reliable classification of calcification and mass type breast abnormalities into malignant and benign classes [36].

Chapter 3

Mammograms and Digital Mammography

3.1 WHAT IS MAMMOGRAM?

A mammogram is an x-ray exam of the breast that's used to detect and evaluate breast changes.

Modern mammography came into action in late 1960's when special x-ray machines were designed and used just for breast imaging whereas x-rays are being used to examine breast tissue since a century ago, by the German surgeon, Albert Salomon, Since 1960's, the technology has advanced a lot and has taken a huge leap, and nowadays mammograms are very different even from those of the 1980s and 1990s.

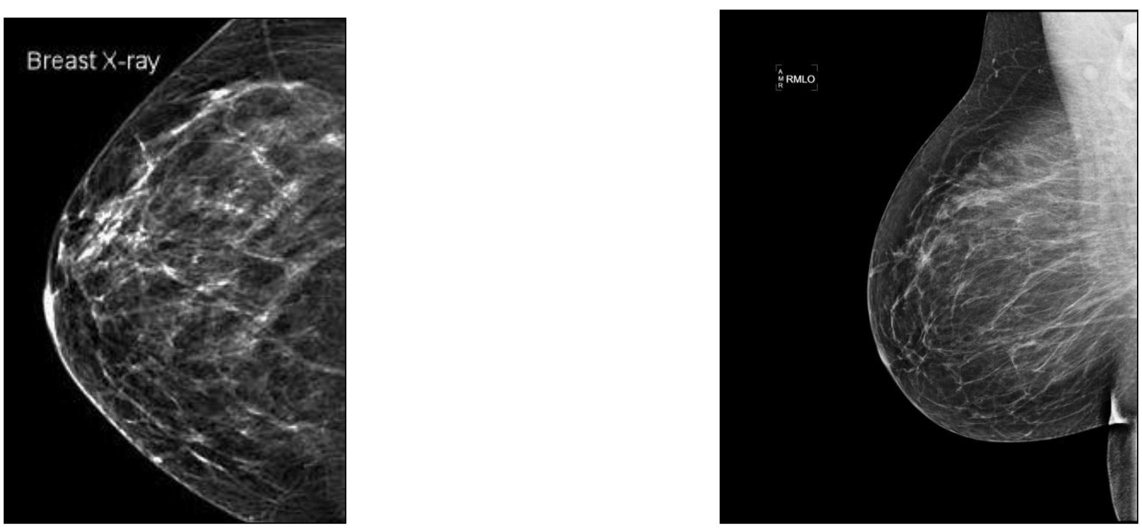

Figure 1 Left image showing x-ray mammogram of breast and right image showing a digital mammogram.

Today, lower energy x-rays machines are being used for acquiring mammograms. These x-rays do not pass through tissue as easily as those used for routine chest x-rays or x-rays of the arms or legs, and this improves the image quality. Mammograms today expose the breast to much less radiation compared with those in the past. A comparison has been as shown in Figure 1[37] [38].

3.2 TYPES OF MAMMOGRAMS

3.2.1 Screening Mammograms look for Signs of Cancer

Screening mammogram are x-ray exams of the breasts that are used for women who have no breast symptoms. The ultimate goal of a screening mammogram is to find breast cancer when it's too small to be felt by a woman or her doctor. Finding small breast cancers early (before they have grown and spread) with a screening mammogram greatly improves a woman's chance for successful treatment.

A screening mammogram usually takes 2 x-ray pictures (views) of each breast. Some women, such as those with large breasts, may need to have more pictures to see as much breast tissue as possible.

3.2.2 Diagnostic Mammograms Investigate Possible Problems

A woman with a breast problem (for instance, a lump or nipple discharge) or an abnormal area found in a screening mammogram typically gets a diagnostic mammogram. It's still an x-ray exam of the breast, but it's done for a different purpose. During a diagnostic mammogram, additional pictures are taken to carefully study the area of concern. In most cases, special pictures are enlarged to make a small area of suspicious breast tissue bigger and easier to evaluate. Other types of x-ray pictures can be done, too, depending on the type of problem and where it is in the breast. A diagnostic mammogram may offer a closer look and show that an area that looked abnormal on a screening mammogram is actually normal. When this happens, the woman goes back to routine yearly screening.

A diagnostic mammogram could also show that an area of abnormal tissue probably is not cancer, but the radiologist may not be ready to say that the area is normal based on these pictures alone. When this happens it's common to ask the woman to return to be rechecked, usually in 4 to 6 months.

The results of the diagnostic work-up may suggest that a biopsy is needed to find out if the abnormal area is cancer. If the doctor recommends a biopsy, it does not mean that one has cancer. About 80% of all breast changes that are biopsied are found to be benign (not cancer). If a biopsy is needed, one should discuss the different types of biopsy with their doctor to decide which type is best suited [39] [40].

3.3 METHODOLOGY OF MAMMOGRAPHY

When a mammogram is done, the breast is concisely compressed or squeezed between 2 plates attached to the mammogram machine—an adjustable plastic plate (on top) and a fixed x-ray plate (on the bottom). The bottom plate holds the x-ray film, or the digital detector that makes the image. The technologist compresses the breast to keep it from moving, and to make the layer of breast tissue thinner. These steps decrease the x-ray exposure and make the picture sharper. Sometimes the compression can feel uneasy or uncomfortable and even painful for some women; it only lasts a few seconds and is needed to get a good picture. Talk to the technologist if one experiences any pain. The doctor can reposition the subject to make the pressure as comfortable as possible. The entire process for a mammogram takes about 20 minutes.

Mammograms produce a black and white x-ray picture of the breast tissue. Depending on the type of machine, the picture is either on a large sheet of film or is an electronic image that can be seen on a computer screen. These two ways of doing a mammogram are much the same. The differences are in the way the picture is recorded, looked at by the doctor, and stored.

The machines that develop the mammogram on x-ray film are screen filled units. Full-field digital mammography units capture the picture in a digital format which can be seen on a computer screen. Most mammogram machines in use today are full-field digital units. Figure 2 shows a full field digital unit.

For the most part, regular screen-film mammograms are as accurate as digital mammograms. But digital mammograms have been shown to have some unique advantages. Some studies have found that women who have questionable areas on their mammogram have to return less often for extra imaging tests because with digital mammograms, the original pictures can be magnified and looked at in many different ways on the computer screen. Several studies have also found that digital mammograms were more accurate in finding cancers in women younger than 50 and in women with dense breast tissue. It’s important to remember that standard film mammograms also work well for these groups of women, and that women should still get their regular mammograms, even if digital mammography is not available.

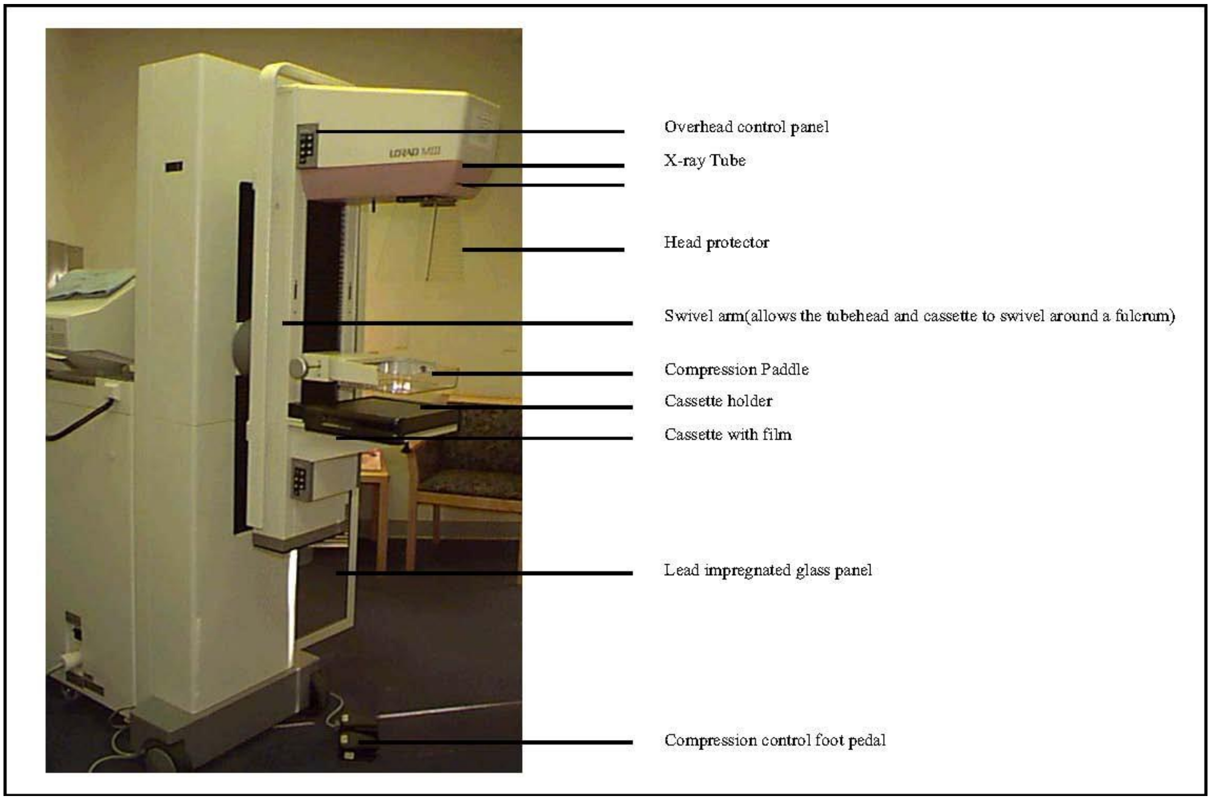

Figure 2 Image showing a mammographic equipment machine.

It does not matter what kind of x-ray image is taken be it film or electronic – it’s interpreted (or “read”) by a doctor, mostly a radiologist. *Radiologists* are the particular doctors who have special training in diagnosing diseases by looking at pictures of the inside of the body produced by x-rays, sound waves, magnetic fields, or other methods. Other doctors who treat breast diseases may look at the mammogram, too.

3.3.1 Reading Mammograms

Reading mammograms is quite challenging. The way the breast looks on a mammogram alters a great deal from woman to woman. And some breast cancers may induce changes in the mammogram that are hard to notice. If one has had mammograms in the past, it’s very important that the radiologist has the latest x-ray films or digital pictures so they can be compared with the new ones. Comparing the pictures helps the doctor find small changes that might have occurred and detect cancer as early as possible. Because it can be hard to get the older pictures, it's best to find a facility that one is comfortable with and plan to get mammograms there each year with regular intervals of time.

3.4 WHAT DOES A MAMMOGRAM DEPICTS?

3.4.1 Breast Density

The report of the mammogram comprise of an assessment of breast density or state that one has dense breasts. Breast density is based on how much part of the breast is made up fatty tissue vs. how much is made up of fibrous and glandular tissue. Dense breasts are not normal, as they are more prone towards the risk of developing a cancer. Although dense breast tissue can make it harder to find cancers on a mammogram, at this time, experts do not agree what other tests, if any, should be done in addition to mammograms in women with dense breasts who aren't in a high-risk group (based on gene mutations, a strong family history of breast cancer, or other factors) [41] [42].

3.4.1.1 BI-RADS reporting for breast density

Mammogram reports can also include an assessment of breast density. BI-RADS classifies breast density into 4 groups:

BI-RADS 1: The breast is almost entirely fat

This means that fibrous and glandular tissue makes up less than 25% of the breast.

BI-RADS 2: There are scattered fibroglandular densities

Fibrous and glandular tissue makes up from 25 to 50% of the breast.

BI-RADS 3: The breast tissue is heterogeneously dense

The breast has more areas of fibrous and glandular tissue (from 51 to 75%) that are found throughout the breast. This can make it hard to see small masses (cysts or tumors).

BI-RADS 4: The breast tissue is extremely dense

The breast is made up of more than 75% fibrous and glandular tissue. This can lead to missing some cancers. In some states, the summary of the mammogram report that is sent to patients must contain information about breast density. This information may be worded in lay language instead of the BIRADS categories. Women whose mammograms show BI-RADS 3 or 4 for breast density may be told that they have “dense breasts.” Figure 3 shows the comparison of BI-RADS category of breast density into 4 groups.

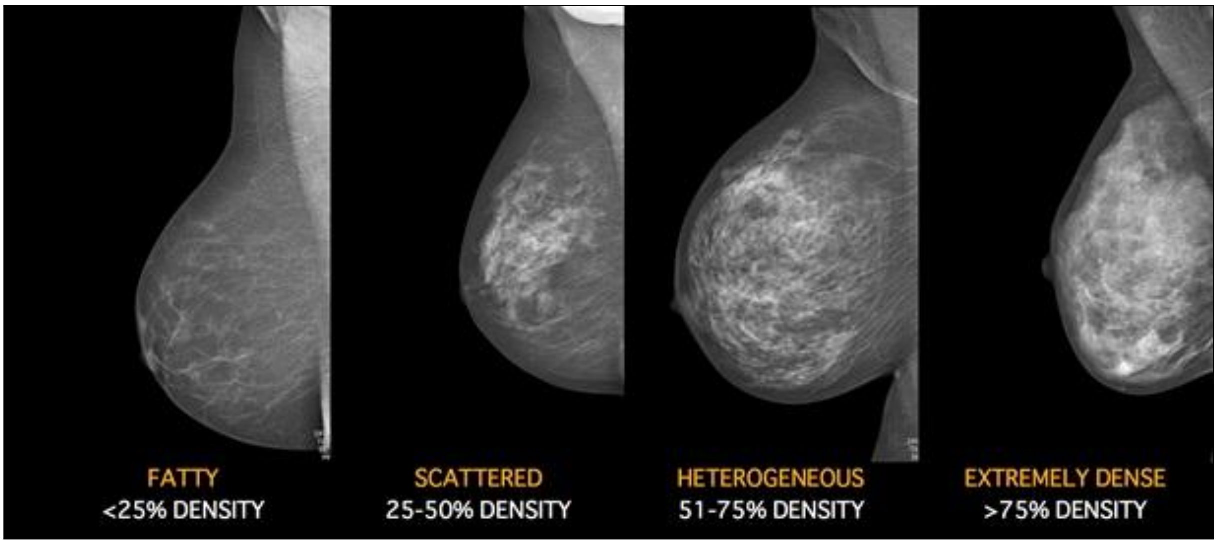

Figure 3 Image showing BIRADS category of breast density.

A mammogram may show something suspicious, but by itself it can’t prove that an abnormal area is cancer. If a mammogram raises a suspicion of cancer, a tissue sample from the suspicious area must be removed and examined under the microscope to find out if it’s cancer. The doctor reading the mammogram will look for different types of changes.

3.4.2 Calcifications

Calcifications are tiny mineral deposits within the breast tissue. They look like small white spots on a mammogram. They may or may not be caused by cancer. There are 2 types of calcifications.

3.4.2.1 Macrocalcifications

Macrocalcifications are coarse (larger) calcium deposits that are probably due to changes in the breasts caused by aging of the breast arteries, old injuries, or inflammation. These deposits do not require any biopsy deposits as these are related to non-cancerous conditions. These are most likely found in about half the women over 50 years and in 1 of 10 women under 50.

3.4.2.2 Microcalcifications

Microcalcifications (shown in Figure 4) are tiny particles of calcium in the breast. They may be seen alone or in clusters. These if seen on a mammogram are of more concern than macrocalcifications, but they do not invariably signify that cancer is present. The radiologist can estimate from the shape and layout of microcalcifications the probability of cancer being present. In most cases, the presence of microcalcifications does not mean a biopsy is needed. But if the microcalcifications have a suspicious look and pattern, a biopsy will be recommended. A biopsy is the only way to tell if cancer is really present [43].

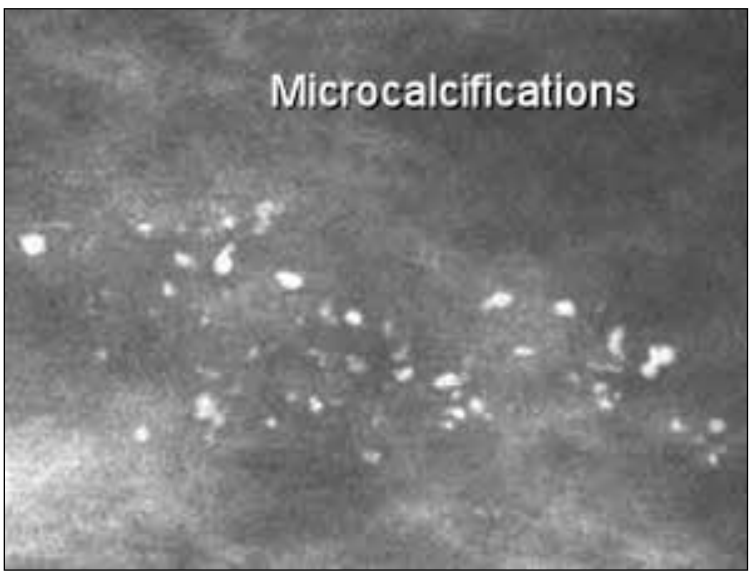

Figure 4 Mammogram showing microcalcifications.

3.4.3 A mass or cyst

A mass (shown in Figure 5), having calcifications or not, is another significant change seen on a mammogram. The areas that look unnatural are masses and they can be lots of things, including cysts (non-cancerous, fluid-filled sacs) and non-cancerous solid tumors (such as fibroadenomas).

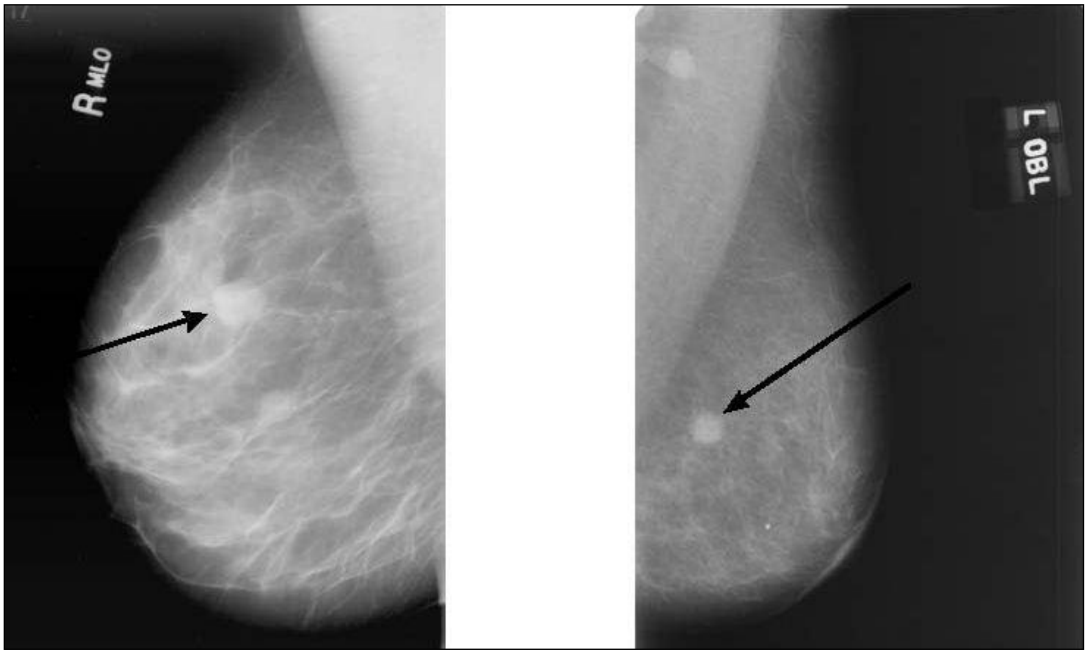

Figure 5 the left mammogram showing benign mass and the right mammogram showing malignant mass.

Cysts can be simple either fluid-filled sacs (known as *simple cysts*) or can be partially solid (known as *complex cysts*). Simple cysts are benign (harmless) and don’t need to be biopsied. Any other type of mass (complex cyst or a solid tumor) might need to be biopsied to be sure it isn’t cancer so that it does not become malignant later on in the future.

A cyst and a tumor can feel the same on a physical exam. They can also look the same on a mammogram. To confirm that a mass is really a cyst, a breast ultrasound is often done. Another option is to draw out the fluid from the cyst with a thin, hollow needle. If a mass is not a simple cyst (partly solid); more imaging tests may be done. Some masses can be watched with regular mammograms, while others may need a biopsy. The size, shape, and margins (edges) of the mass may help the radiologist determine if cancer is likely to be present. Having some prior mammograms available for the radiologist is very important. They can help show that a mass or calcification has not changed for many years into something malignant. This would mean that it's likely not cancer and a biopsy is not needed [44]. Figure 6 shows a mammogram which is showing a cyst.

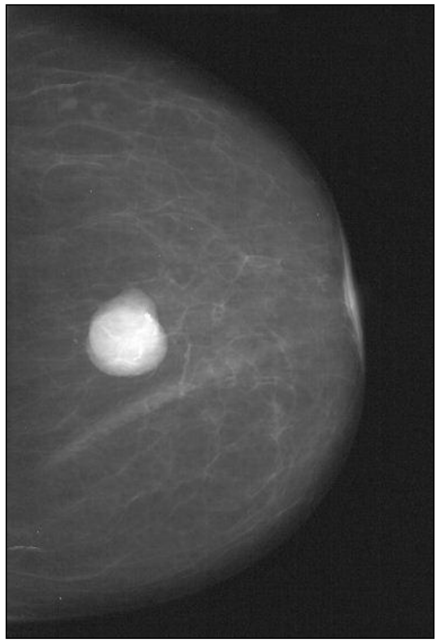

Figure 6 Mammogram showing a cyst.

3.5 LIMITATIONS OF MAMMOGRAMS

As is the case with most medical tests, mammography has limitations. Although breast cancer screening is the best way nowadays to find cancer early, finding cancer early does not always reduce a woman's chance of dying from breast cancer. Even though mammograms can detect breast cancers too small to be felt, treating a small tumor does not always mean it can be cured. A fast-growing or aggressive cancer may have already spread before it's found. The value of a screening mammogram also depends on a woman's overall health status. Detecting breast cancer early may not help prolong the life of a woman who has other kinds of serious or life-threatening health problem such as congestive heart failure, end stage renal disease, or chronic obstructive pulmonary (lung) disease.

3.5.1 False-negative results

A false-negative mammogram appears normal even though breast cancer is present. Overall, screening mammograms miss about 1 in 5 breast cancers. The main cause of false-negative results is high breast density. False negatives occur more often among younger women than among older women because younger women are more likely to have dense breasts. Breasts usually become less dense as women age. False-negative results can delay treatment and promote a false sense of security for the woman.

3.5.2 False-positive results

A false-positive mammogram looks abnormal but no cancer is actually present. Abnormal mammograms require extra testing (diagnostic mammograms, ultrasound, and sometimes biopsy) to find out if cancer is present. False-positive results are more common in women who are younger, have dense breasts, have had breast biopsies, have breast cancer in the family, or are taking estrogen. With annual screening, over a 10-year period the odds that a woman will have a false-positive finding are greater than 50%. The odds of a false-positive finding are highest for the first mammogram, and are lower on subsequent mammograms. Women who have prior films available for comparison reduce the odds of a false-positive finding by 50%. False-positive mammograms can cause temporary anxiety. The extra tests needed to be sure cancer isn't there cost time and money and also cause physical discomfort. Still, most studies of attitudes towards false positives have shown that women accept false positive findings as part of the process of finding breast cancer early [45].

3.5.3 Radiation exposure

Mammograms require very small doses of radiation. The risk of harm from this radiation is extremely low, but in theory, repeated x-rays might have the potential to cause cancer. Still, the benefits of mammography outweigh any possible harm from the radiation exposure. Women should always let their health care providers and x-ray technologists know if there is any chance that they are pregnant, because radiation can harm a growing fetus.

3.6 IMPROVING MAMMOGRAMS

Although a mammogram is an excellent way to find most breast cancers when they are small and most curable, it does not detect all breast cancers. Newer techniques are being looked at to try to make mammograms more accurate.

3.6.1 Computer-aided detection and diagnosis

Computer-aided detection and diagnosis (CAD) was developed to help radiologists find suspicious changes on mammograms. This technology can be used with standard film mammograms or with digital mammograms.

Computers can help doctors find abnormal areas on a mammogram by acting as a second set of eyes. For standard mammograms, the film is fed into a machine which converts the image into a digital signal that is then analyzed by the computer. The technology can also be applied to an image captured with a digital mammogram. The computer then displays the picture on a video screen, with markers pointing to areas the radiologist should check more closely.

Early research on CAD showed a clear improvement in finding small cancers, with only a small increase in the number of women who had to come back for more tests. But studies of CAD in community practice have shown mixed results. Some showed a clear benefit from the use of CAD, and others showed that it did not find more cancers or find cancers earlier, but did increase the number of women who needed to come back for more tests and/or to have breast biopsies. Current research suggests that CAD is not a substitute for experience and expertise in reading mammograms. In other words, CAD is only helpful when the radiologists are experienced and have expertise in reading mammograms [46].

3.6.2 Tomosynthesis (3D mammography)

This technology is basically an extension of a digital mammogram. For this test, the breast is compressed once and a machine takes many low-dose x-rays as it moves over the breast. The images can then be combined into a 3-dimensional picture. Although this uses more radiation than most standard 2-view mammograms, the dose still is below the maximum dose allowed by the Mammography Quality Standards Act. This technology may allow doctors to see problem areas more clearly, which can mean fewer patients will need to be called back for more tests.

A breast tomosynthesis machine shown in Figure 7 was approved by the Food and Drug Administration (FDA) in 2011 for use in the United States, but the role of this technology in screening and diagnosis is still not clear. Not all health insurance covers tomosynthesis, so one may want to check with the insurance company if this is covered by them. Figure 8 shows the comparison between a traditional x-ray mammogram and a digital breast tomosynthesis. One can see that the tomosynthesis digital mammogram is much clear than the traditional mammogram and shows the cancerous tissue with much depth and clarity.

Figure 7 Image showing equipment for tomosynthesis.

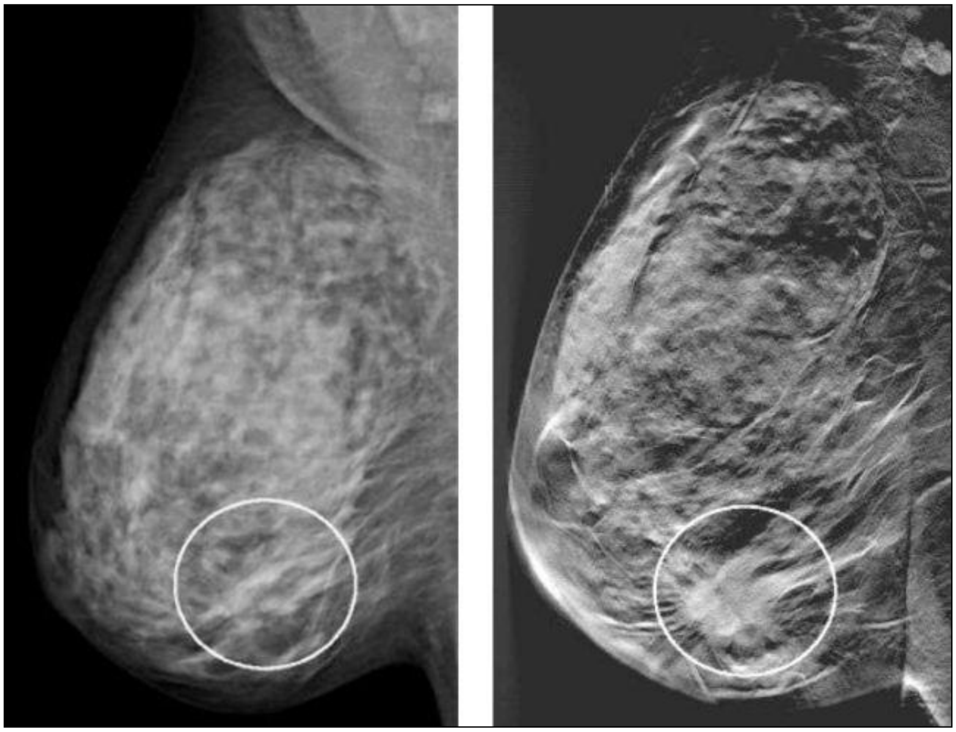

Figure 8 Left image showing tradition x-ray mammogram and right image showing digital breast tomosynthesis.

Chapter 4

Texture Analysis

4.1 INTRODUCTION

Many real objects in nature do not possess uniformity in their surface structure or the way they reflect light. Instead they show some repetitive pattern. In broad this repetitive pattern can be defined as texture. In other way texture can be understood as uniformity in being non-uniform. But there is no clear definition of texture. Many pioneers in different fields have defined the word texture in different ways applicable to their field and work. Some of the definitions which can give broad understanding of texture are given below.

- "We may regard texture as what constitutes a macroscopic region. Its structure is simply attributed to the repetitive patterns in which elements or primitives are arranged according to a placement rule." [47]
- "A region in an image has a constant texture if a set of local statistics or other local properties of the picture function are constant, slowly varying, or approximately periodic."[48]
- "The notion of texture appears to depend upon three ingredients: (i) some local 'order' is repeated over a region which is large in comparison to the order's size, (ii) the order consists in the non-random arrangement of elementary parts, and (iii) the parts are roughly uniform entities having approximately the same dimensions everywhere within the textured region."

For a layman eye is a god gifted instrument which sees everything at a glance. But in-fact our visual system does not 'sees' a scene; it tries to understand and compare different objects in the 'scene'. The process of seeing a scene is so much complicated that only a person working on brain or machine vision can understand the complexity and process behind it.

Normally when one sees a scene one just looks at the global variations in it. A person does not really look at the local variations and the reason behind why different scenes and objects really look different i.e. texture. Figure 9 shows some of the natural texture pictures and Figure 10 shows some of the synthetic texture pictures which can help one understand the meaning of texture and importance of local and global patterns which actually differentiate different objects in a scene or view.

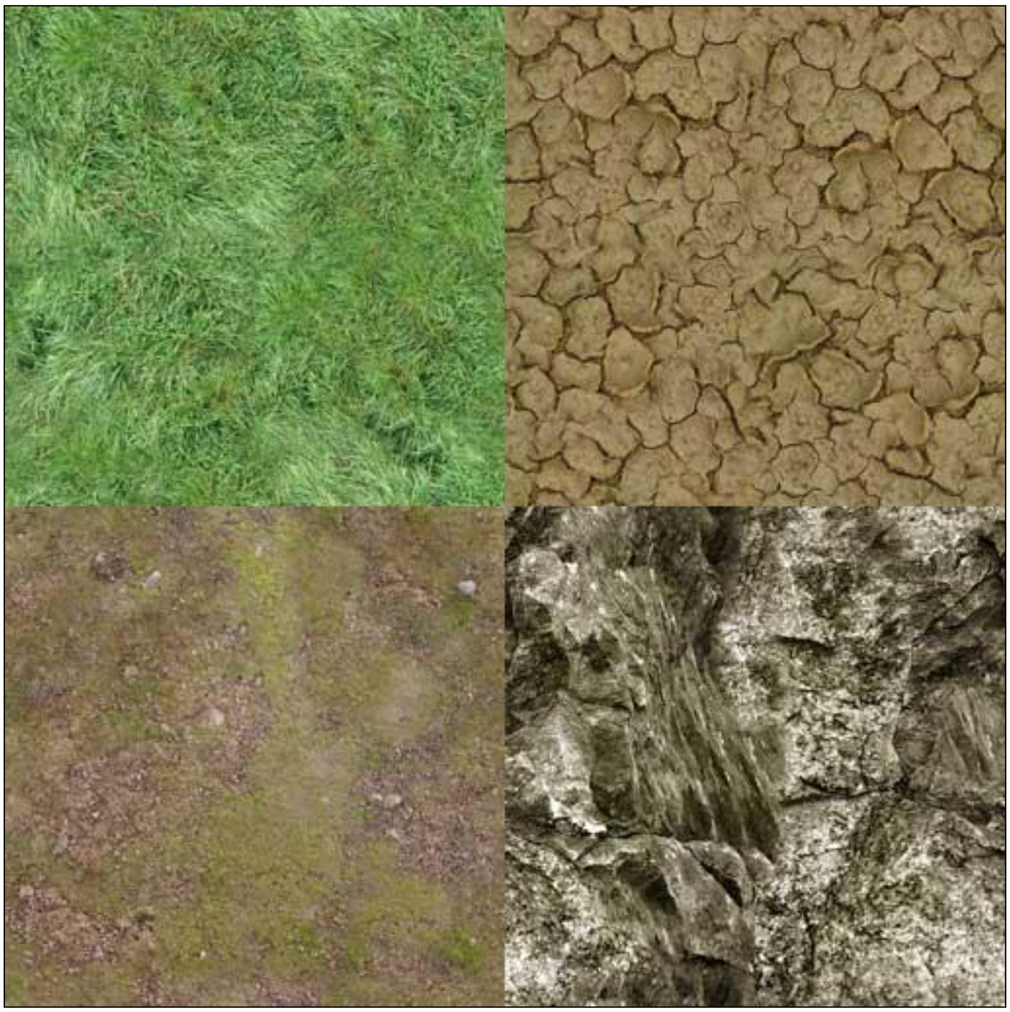

Figure 9 Image showing some natural textures.

Figure 10 Image showing some synthetic textures.

4.2 TEXTURE ANALYSIS: A STATISTICAL APPROACH

For any given image, statistical methods examine the spatial distribution of gray values and this is done by calculating the local features at each point, thus deducing a set of statistics from the statistical distributions of the local features. Statistical approach for texture analysis (statistical methods) can be divided into first order (one pixel), second-order (two pixels) and higher-order (three or more pixels) statistics based on the number of pixels which define the local feature utilized to calculate the spatial distribution of gray values [49].This is because of the fact that the spatial distribution of gray values is one of the defining and determining qualities/characters of texture. The basic difference is that first-order statistics calculate properties such as average and variance of individual pixel values, neglecting the spatial interaction between image pixels, on the other hand second and higher order statistics calculate properties of two or more pixel values occurring at particular positions relative to each other.

Statistical methods concede characterizations of textures as fine, coarse etc. Thus one measure of texture is based on the primitive size, which could be the average area of these primitives of relatively constant gray level. The average could be taken over some set of primitives to measure its texture or the average could be about any pixel in the image. If the average is considered inside a primitive centered at each pixel in the image, the result can be utilized to develop a texture image in which a large gray level at a pixel points, for example, that the average primitive size is large in a region around that pixel. The average shape measure of these primitives, such as P^2/A, where P is the perimeter and A is the area of the primitive could also be used as texture measure.

4.3 GLCM

Spatial gray level co-occurrence estimates image properties related to second-order statistics which considers the relationship among pixels or groups of pixels (usually two). Haralick [50] suggested the use of gray level co-occurrence matrices (GLCM) which have become one of the most well-known and widely used texture features. This method is based on the joint probability distributions of pairs of pixels. GLCM show how often each gray level occurs at a pixel located at a fixed geometric position relative to each other pixel, as a function of the gray level. The (1,3) entry in a matrix for right neighbors, for example, would show the frequency or probability of finding gray level 3 immediately to the right of pixel with gray level 1.

i	j

Figure 11 Sample of an image.

1	3	2
3	1	3
2	1	4

Figure 12 Original Image.

	0	**1**	**2**	**3**
0	0	0	2	1
1	1	0	0	0
2	1	0	0	0
3	0	1	0	0

Figure 13 Co-occurrence matrix.

Figure 11 shows a sample of an image. Figure 12 shows a 3 x 3 image and its 4 gray level co-occurrence matrices are presented in the Figure 13. The number of threshold levels is 4. The 2 in the co-occurrence matrix indicates that there are two occurrences of a pixel with gray level 3 immediately to the right of pixel with gray level 1.

The size of co-occurrence matrix will be the number of threshold levels. When we consider neighboring pixels, the distance between the pair of pixels is 1. However, each different relative position between the two pixels to be compared creates a different co-occurrence matrix. If the edges between the neighboring elements (texels) are slightly blurred, nearby neighbors may be very similar in gray level, even near the edges of the texels. In such cases it will be better to base the co-occurrence matrix on more distant neighbors. For example, the matrix entry mij could represent the number of times gray level j was found 3 pixels to the right of gray level i in the region.

Rather than using gray level co-occurrence matrix directly to measure the textures of images and regions, the matrices can be converted into simpler scalar measures of texture. For example, in an image where the gray level varies gradually, most of the non-zero entries for the right neighbors will be near the main diagonal because the gray levels of neighboring pixels will be nearly equal. A way of quantifying the lack of smoothness in an image is to measure the weighted average absolute distance d of the matrix entries from the diagonal of the matrix.

$$d = \frac{1}{M} \sum_{i,j} |i-j|\, m_{ij} \qquad \text{(Equation 4.1)}$$

where: $M = \sum_{ij} m_{ij}$

The example in Figure 14 and Figure 15 is for a gradually changing vertical edge and hence all the non-zero entries tend to concentrate towards the diagonal with the result |i-j| = 0 for each entry and d=0 for all the entries.

0	1	2	3	4

Figure 14 Original Image with graduation in gray level.

	0	1	2	3	4
0	0	1	0	0	0
1	0	0	1	0	0
2	0	0	0	1	0
3	0	0	0	0	1
4	0	0	0	0	0

Figure 15 Co-occurrence matrix with non-zero entries concentrated near the main diagonal.

Where i, j are the size (no. of rows and columns) of the co-occurrence matrix and d is the absolute distance of the matrix entries from the diagonal of the matrix. If the neighbors tend to have very different gray levels, most of the entries will be far from the diagonal and the value of d will be large, which indicates an uneven edge which can be used to represent a terrain [51] [52]. Figure 16 shows some gray-level textures.

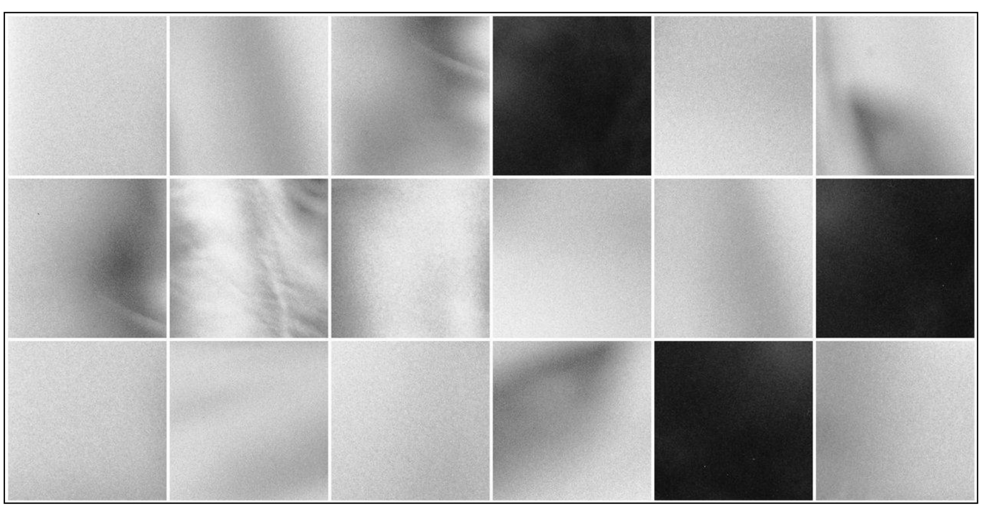

Figure 16 Image showing some gray level textures.

Chapter 5

Material and Methodology

5.1 DATABASE

The Mammographic Image Analysis Society (MIAS) [53] is an organization of UK research groups interested in the understanding of mammograms and has generated a database of digital mammograms. Films taken from the UK National Breast Screening Programme have been digitized to 50 micron pixel edge with a Joyce-Loebl scanning microdensitometer, a device linear in the optical density range 0-3.2 and representing each pixel with an 8-bit word. The database contains 322 MLO digitized mammograms corresponding to left and right breast of 161 women and is available on 2.3GB 8mm (Exabyte) tape. It also includes radiologist's "truth"-markings on the locations of any abnormalities that may be present. The database has been reduced to a 200 micron pixel edge and padded/clipped so that all the images are 1024 x1024. Mammographic images are available via the Pilot European Image Processing Archive (PEIPA) at the University of Essex. Each pixel was described as a 8-bit word. The database contains mammographic images named as “mdbXXXBS”, where:

- “XXX” represent the number of the image ranging from 001 to 322.
- “B” is the side of breast: “l” or “r” (left or right).
- “S” is the size of the image, which can be ”s” for small images (1600 x 430 pixels),

”m” for medium size images (2048 x 4320 pixels), ”l” for large images (2600 x 4320 pixels), and ”x” for extra-large images (4000 x 5200 pixels). In our master thesis the medium size images “m” has been used.

5.2 METHODOLOGY

The algorithm developed has been applied on the dataset is as follows:

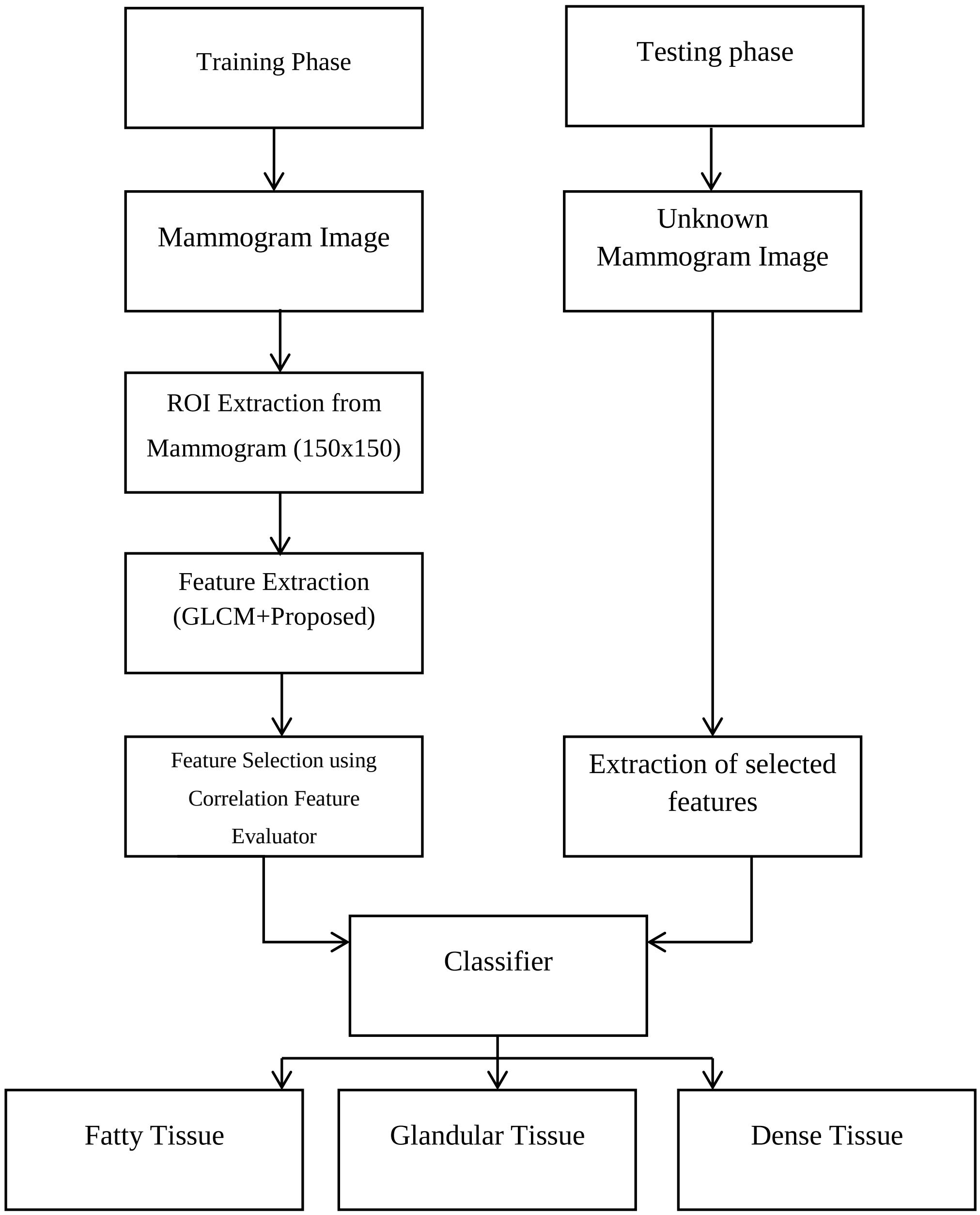

Figure 17 Flowchart showing the methodology to be followed for the classification of the mammograms.

5.2.1 Pre-Processing

In the first step of the proposed model, region-of-interest (ROI) of 150 by 150 pixels is extracted from each mammogram used in the algorithm from the area right behind the nipple (as shown in Figure 18) because approaches based on texture feature analysis typically analyse the instinct characteristics of texture features extracted from regions of interest (ROIs) to classify ROIs into well-known knowledge categories. Analysis of texture centred in mammograms is necessary because it is useful in order to identify specific types of regions of interest. Further if we take into account that digital mammography produces high resolution gray level images where textures play an important role, feature extraction descriptors are needed with the purpose to identify and select a set of distinguishing and sufficient features to characterize a mammogram. In fact, it is one way to achieve texture information with basic use to classify the image data. This information contains data about spatial distribution on the intensity pixels within defined regions in a grey scale image. Thus, the texture of a region describes the pattern of spatial variation of grey tones in a neighbourhood.

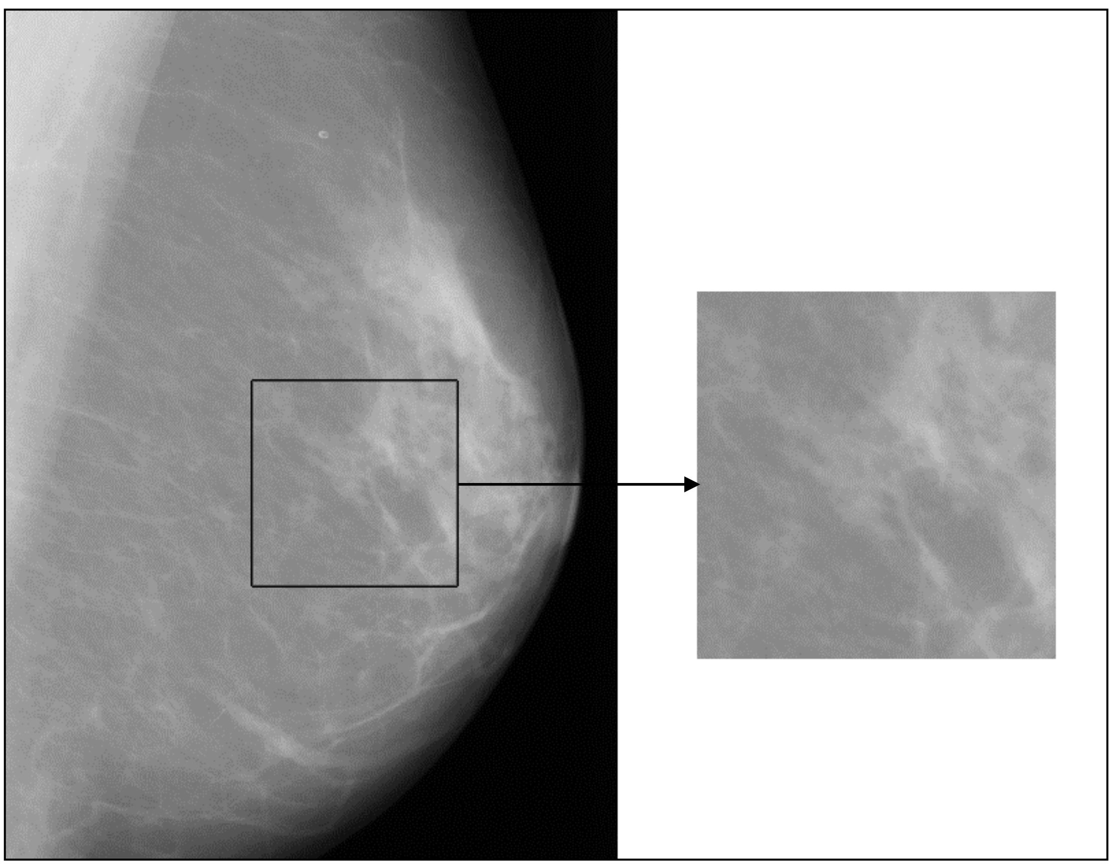

Figure 18 Example of an ROI (150 x 150) extracted from a mammogram “mdb076”.

5.2.2 GLCM Features

In the second step GLCM texture features [54] are extracted from each ROI. Basically the idea is to achieve highest possible classification rate and to find out the best possible features. The features are extracted from the gray-level co-occurrence matrix. It is a statistical method that considers the spatial relationship of pixels, also known as the gray-level spatial dependence matrix [55]. By default, the spatial relationship is defined as the pixel of interest and the pixel to its immediate right (horizontally adjacent), but one can specify other spatial relationships between the two pixels. In this study [1,0] offset has been used. Each element (I, J) in the resultant GLCM is simply the sum of the number of times that the pixel with value I occurred in the specified spatial relationship to a pixel with value J in the input image. $p(i,j)$ is the (i,j)th entry in a normalized GLCM, $px(i)$ is the ith entry in the marginal probability matrix obtained by summing the rows of $p(i,j)$, Ng is the number of distinct gray levels in GLCM, HX and HY are entropies of px and py. The feature extraction has been done only for the angle of 0º and distance between the pixel of interest and its neighbour equal to 1. Table 1 shows some of the extracted features and below are some GLCM features and their significance:

Contrast: A measure of the intensity contrast between a pixel and its neighbor over the whole image. Contrast is 0 for a constant image.

Correlation: A measure of how correlated a pixel is to its neighbor over the entire image. Correlation is `NaN` for a constant image.

Energy: A measure of uniformity in the range [0,1]. Uniformity is 1 for a contrast image.

Entropy: Measures the randomness of the image. The entropy is 0 when all *Pij* 's are 0 and is maximum when all *Pij* 's are equal.

Homogeneity: A value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Homogeneity is 1 for a diagonal GLCM.

Variance: The expected value of the square of the deviations of a random variable from its mean value.

Table 1 GLCM Features and their Mathematical Formulae.

Feature No.	GLCM Features	Mathematical Formula
1.	Contrast	$f_1=\sum_{n=0}^{N_g-1} n^2\{\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} p(i,j)\vert i-j\vert = n\}$
2	Correlation	$f_2 = \frac{\sum_i\sum_j (ij)p(i,j) - \mu_x\ \mu_y}{\sigma_x\,\sigma_y}$
3	Energy	$f_3 = \sum_i\sum_j p(i,j)^2$
4	Entropy	$f_4 = -\sum_i\sum_j p(i,j)\log(p(i,j))$
5	Homogeneity	$f_5 = \sum_i\sum_j \frac{1}{1+(i-j)^2}\ p(i,j)$
6	Variance	$f_6 = \sum_i\sum_j (i-j)^2\ p(i,j)$
7	Sum Average	$f_7 = \sum_{i=2}^{2N_g} i p_{x+y}(\mathrm{i})$
8	Sum Variance	$f_8 = \sum_{i=2}^{2N_g} (i - f_{14})^2 p_{x+y}(\mathrm{i})$
9	Sum Entropy	$f_9 = -\sum_{i=2}^{2N_g} p_{x+y}(\mathrm{i})\ log(p_{x+y}(i))$
10	Difference Variance	$f_{10} = \sum_{i=0}^{N_g-1} (i - f_6)^2\, p_{x-y}\,(i)$
11	Difference Entropy	$f_{11} = \sum_{n=0}^{N_g-1} p_{x-y}\,(i)\log(p_{x-y}(i))$
12	Information measure of correlation 1	$f_{12} = \frac{\sum_i\sum_j p(i,j)\,(\log(p(i,j)) - \log(p_x(i)p_y(i)))}{\max(HX,HY)}$
13	Information measure of correlation 2	$f_{13} = \sqrt{1 - e^{-2(a-b)}}$ $a = -\sum_i\sum_j p_x\,(i)p_y(i)\log(p_x(i)p_y(i))$ $b = -\sum_i\sum_j p(i,j)\log(p_x(i)p_y(i))$

5.2.3 Proposed Features

In this thesis, three new features have been proposed:

I. White area
II. Pixel Count
III. Contrast

The three features have been explained below:

5.2.3.1 White Area

All the images are converted into binary images and above certain intensity the white areas of the image are calculated. For example Figure 19 shows an extracted ROI of an image "mdb076" and its binary image in which pixels having pixel intensity equal to or less than 160 were converted to zero intensity (i.e. black region) and pixels having intensity greater than 160 were converted to pixel intensity 255 (i.e. white region). In this feature the area of the white region of all the images are calculated in a certain ranges of intensity. The various intensity ranges taken for the experiment are: White area above intensity 175; white area above intensity 190; white area above intensity 200; white area above intensity 210.

5.2.3.2 Pixel Count

Here, again all the images are converted to binary images but here instead of calculating the area of the white region we counted the number of pixels in the white region. There was a lot of difference in the pixel count of white region for the different ranges of intensity this feature was included for the classification. The several ranges taken for the experiment are: pixel count between the intensity range 110-160; pixel count between the intensity range 140-155; pixel count between the intensity range 180-230; pixel count between the intensity range 190-205; pixel count above intensity 175.

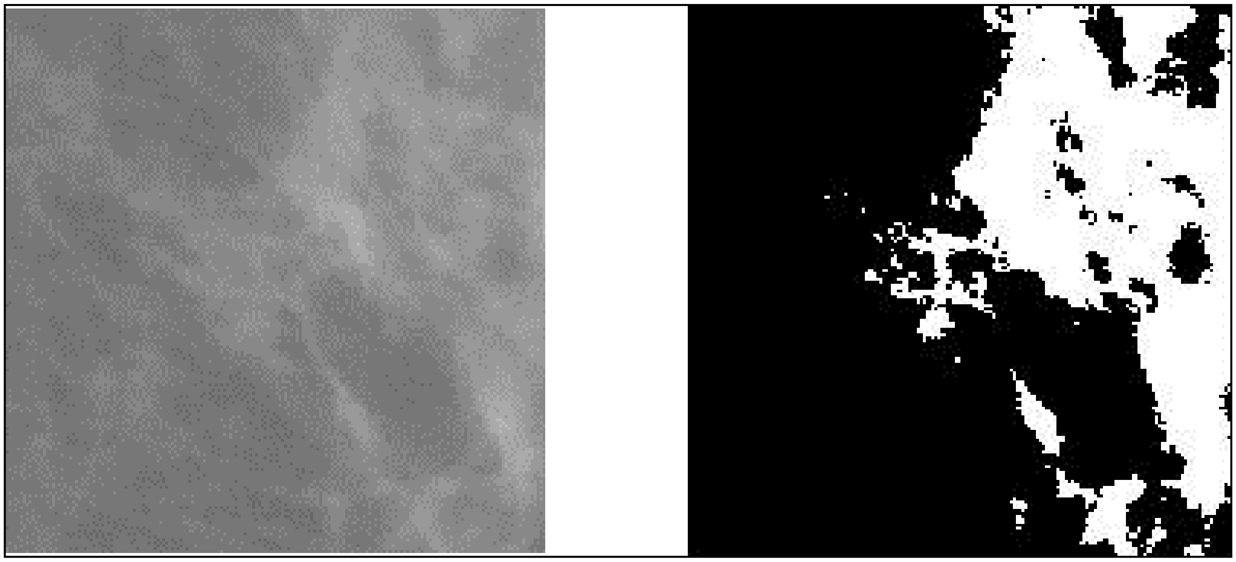

Figure 19 Extracted ROI of “mdb076” (left) and Binary image of extracted ROI of “mdb076” (right).

5.2.3.3 Contrast

The contrast for all the images is calculated manually by calculating the difference between the maximum intensity and minimum intensity of the all the images. The proposed method had two contrasts: contrast in the GLCM features and the contrast in the proposed features. To avoid the contradiction between both, the proposed contrast was normalised between the valures 0 and 1; and its correlation was found out with GLCM contrast. The correlation came out to be 0.096, which was very low, so both the contrasts were used for further feature selection process. So, in all three new intensity based features are proposed and used in this work.

5.3 PATTERN RECOGNITION

Pattern can be defined as a structural or quantitative description of a subject of interest. The subject may be a visible object or a system of data. Exploring and analysis of the pattern hence visualized is called *pattern recognition.* A set of patterns which share some properties in common are defined as *pattern class*. For example, we can easily differentiate between trees and animals because they show some distinct properties in common. Hence trees and animals belong to different class.

In a broad sense, we all have been using pattern recognition techniques in our daily life, although without realizing, that is god gifted to us. Every time we see something we try to differentiate objects of similar properties from others without even feeling the complexity behind it. In machine vision field, there are two types of items for recognition:

1) Recognition of concrete items: There are the items that can be visualized easily. Some of them are spatial and temporal properties of objects. For example they are characters of different language, roads, buildings, cars of different colors etc.

2) Recognition of abstract items: It is mainly a conceptual recognition. For example deciding which car is better to purchase, which book is better to refer. This cannot be decided just by some mathematical expression. It depends on the intellectuality of the person and his/her taste. They are totally conceptual.

For pattern recognition there are two basic factors which are required for the recognition and a few basic steps necessary for making the analysis shown in the Figure 20. The two factors which help in the process of classification, clustering or other application are:

1) Features: a prominent attribute or aspect

2) Classifiers: mechanism or method to define a pattern

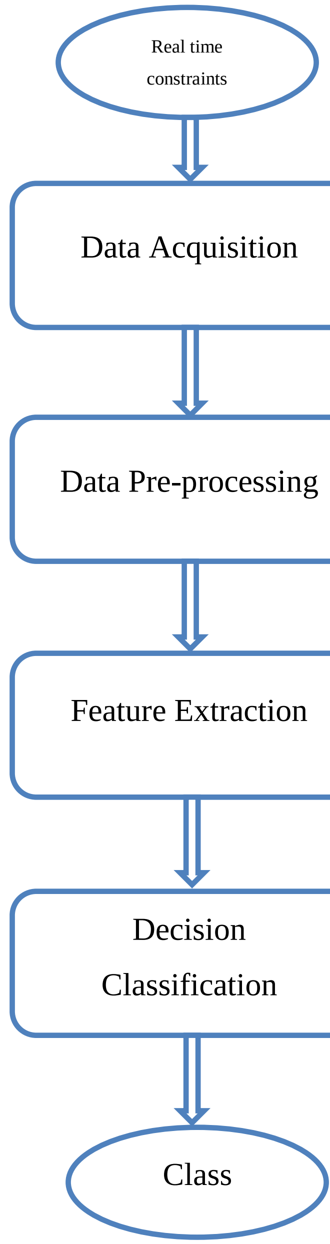

Figure 20 A Conceptual Pattern Recognition System.

5.4 DIMENSIONALITY REDUCTION

Usually, because of availability of large number of measures and the approach of defining a pattern, the pattern space is of high dimensional. But it does not mean every measure or approach of defining a particular 'object' is useful; sometimes some measures simply reflect some other measure and hence there is overflow of data measured. For considering only the principal measures which can best represent the 'object', there is a need of dimensionality reduction technique.

5.4.1 Feature Selection

Feature subset selection is the procedure of keying out and taking out as much irrelevant, repeated and redundant information as possible. It has long been a research area within pattern recognition. This brings down the dimensionality of the data and may allow learning algorithms to operate faster and more effectively which otherwise would be complex and hectic. In some cases, accuracy on future classification can be improved; in others, the result is a more compact, easily interpreted representation of the target concept.

Genari *et al.* state that

"Features are relevant if their values vary systematically with category membership."

In other words, a feature is useful if it is correlated with or predictive of the class; otherwise it is irrelevant [56].

Kohavi and John formalize this definition as:

A feature V_i is said to be relevant *iff* there exists some v_i and c for which $p(V_i = v_i) > 0$ such that

$$p(C = c|V_i = v_i) \neq p(C = c)$$

A feature is said to be redundant if one or more of the other features are highly correlated with it. "Features" are individual tests which measure traits related to the variable of interest i.e. (class) [57]. If the correlation between each of the components in a test and the outside variable is known, and the inter-correlation between each pair of components is given, then the correlation between a composite test consisting of the summed components and the outside variable can be predicted from:

$$r_{zc} = \frac{k r_{zi}}{\sqrt{k + k(k-1) r_{ii}}} \qquad \text{(Equation 5.1)}$$

where r_{zc} is the correlation between the summed components and the outside variable, k is the number of components, r_{zi} is the average of the correlations between the components and the outside variable, and r_{ii} is the average inter-correlation between components.

Equation 5.1 is, in fact, Pearson's correlation coefficient, where all variables have been standardized. It shows that the correlation between a composite and an outside variable is a function of the number of component variables in the composite and the magnitude of the inter-correlations among them, together with the magnitude of the correlations between the components and the outside variable.

5.4.1.1 Correlation Feature Selection

Correlation Feature Selection (CFS) is an easy going filter algorithm that ranks feature subsets according to a correlation based heuristic evaluation function. The result of the evaluation function comprise of the subsets that contain features that are highly correlated with the class and uncorrelated with each other. Irrelevant features should be neglected because they will have low correlation with the class. Redundant features should be screened out as they will be highly correlated with one or more of the remaining features. The feature will be accepted depending on the extent to which it predicts classes in areas of the instance space not already predicted by other features. CFS's feature subset evaluation function (Equation 5.1) is repeated here (with slightly modified notation) for ease of reference:

$$M_s = \frac{kr_{cf}}{\sqrt{k+k(k-1)r_{ff}}} \qquad \text{(Equation 5.2)}$$

where M_S represents heuristic "merit" of a feature subset S containing k features, r_{cf} is the mean feature-class correlation (f ϵ S), and r_{ff} is the average feature-feature inter-correlation. The numerator of equation 5.2 can be thought of as providing an indication of how predictive of the class a set of features are; the denominator of how much redundancy there is among the features.

Equation 5.2 forms the core of CFS and imposes a ranking on feature subsets in the search space of all possible feature subsets. The implementation of CFS used in the experiments allows the user to choose from three heuristic search strategies: forward selection, backward elimination, and best first. Forward selection begins with no features and greedily adds one feature at a time until no possible single feature addition results in a higher evaluation. Backward elimination begins with the full feature set and greedily removes one feature at a time as long as the evaluation does not degrade. Best first algorithm can start with either no features or all features. In the former, the search progresses forward through the search space adding single features; in the latter the search moves backward through the search space deleting single features. To prevent the best first search from exploring the entire feature subset search space, a stopping criterion is imposed. The search will terminate if five consecutive fully expanded subsets show no improvement over the current best subset.

5.4.1.2 Heuristic Search

Searching the space of feature subsets within reasonable time constraints is necessary if a feature selection algorithm is to operate on data with a large number of features. One simple search strategy, called greedy hill climbing, considers local changes to the current feature subset. Often, a local change is simply the addition or deletion of a single feature from the subset. When the algorithm considers only additions to the feature subset it is known as *forward selection*; considering only deletions is known as *backward elimination.*

An alternative approach, called stepwise bi-directional search, uses both addition and deletion. Within each of these variations, the search algorithm may consider all possible local changes to the current subset and then select the best, or may simply choose the first change that improves the merit of the current feature subset. In either case, once a change is accepted, it is never reconsidered. If scanned from top to bottom, the diagram shows all local additions to each node; if scanned from bottom to top, the diagram shows all possible local deletions from each node.

Best first search is an AI search strategy that allows backtracking along the search path. Like greedy hill climbing, best first moves through the search space by making local changes to the current feature subset. However, unlike hill climbing, if the path being explored begins to look less promising, the best first search can back-track to a more promising previous subset and continue the search from there. Given enough time, a best first search will explore the entire search space, so it is common to use a stopping criterion. Normally this involves limiting the number of fully expanded[3] subsets that result in no improvement.

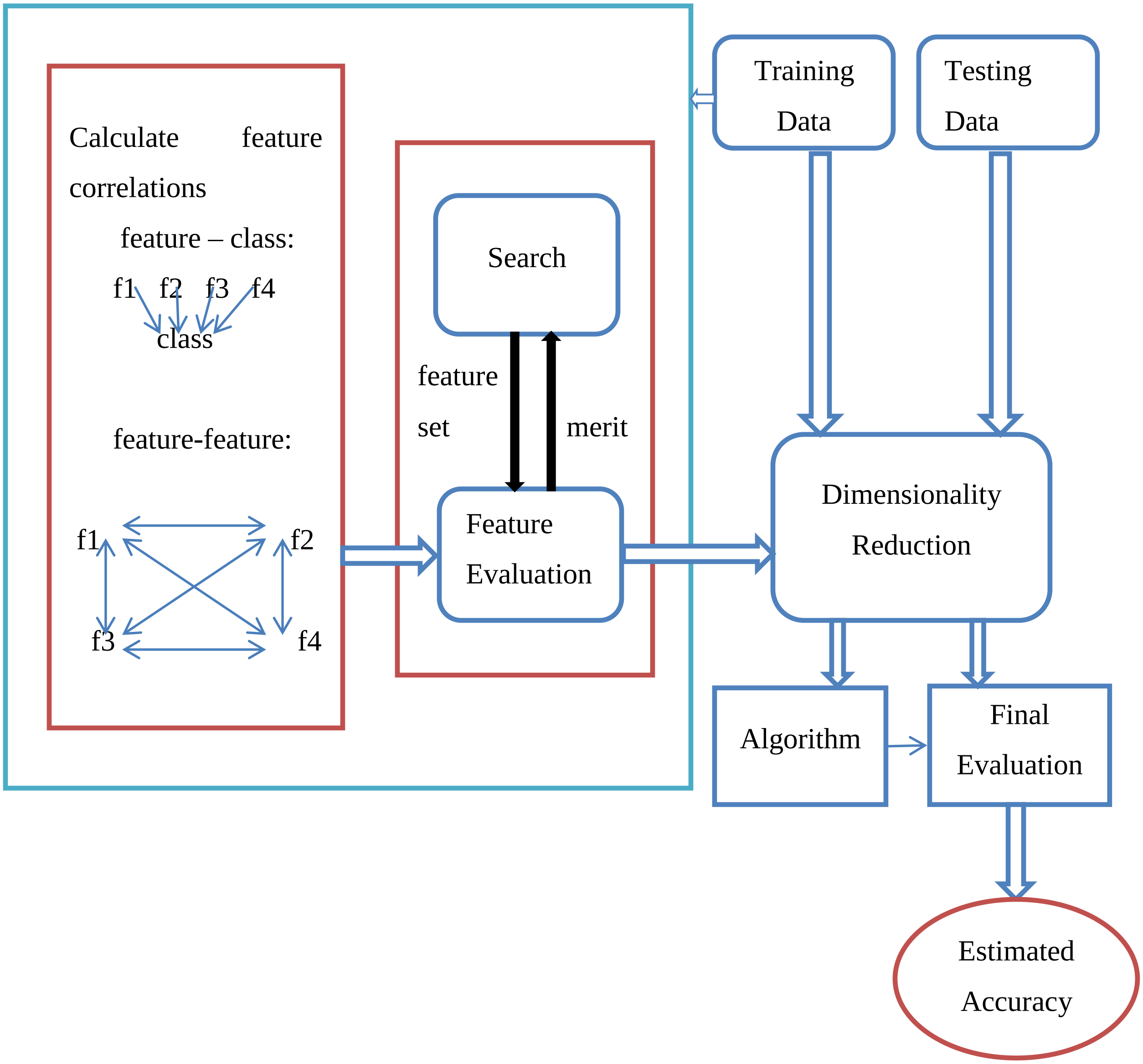

Figure 21 Flowchart showing working of Correlation Feature Selection.

The working of CFS (Correlation-based Feature Selection) is shown in Figure 21. CFS assumes that useful feature subsets contain features that are predictive of the class but uncorrelated with one another. CFS computes a heuristic measure of the "merit" of a feature subset from pair-wise feature correlations and a formula adapted from test theory. Heuristic search is used to traverse the space of feature subsets in reasonable time; the subset with the highest merit found during the search is reported [58].

5.5 CLASSIFICATION

Classification is a task of training a classification model that maps each attribute set x to one of the predefined class labels y. The Figure 22 shows a general approach of a classification problem. First, a training set consisting of records whose class labels are known are provided. The training set is then used to train the classification model, which is eventually used to label the testing set.

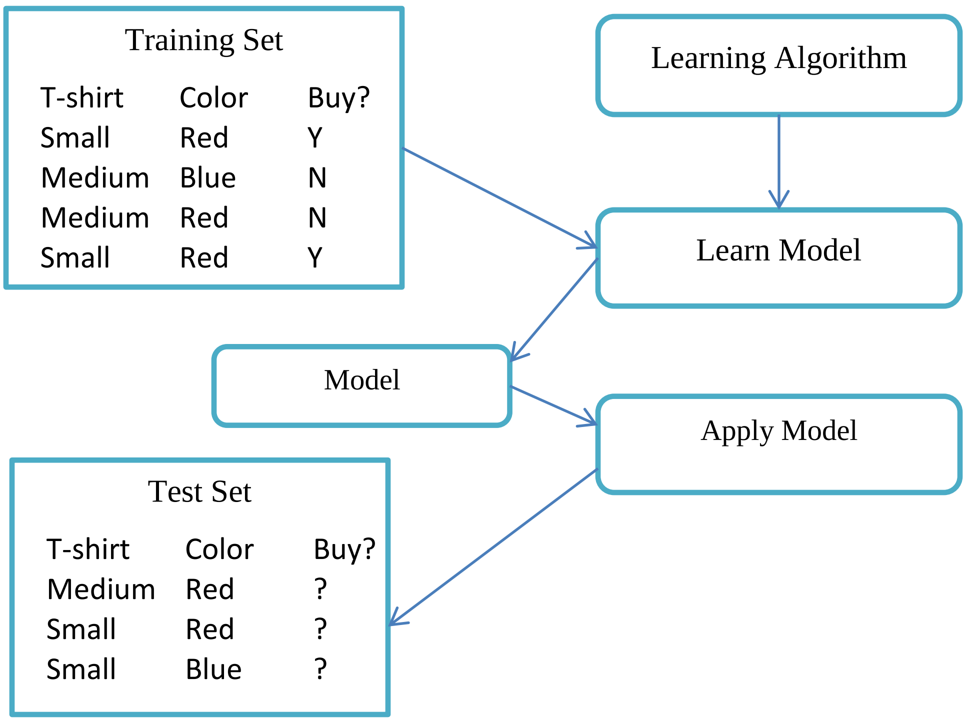

Figure 22 Figure showing a general classification approach.

There is no single classification approach which can be used for every problem. A classification model is to be chosen is problem specific and accuracy of the result differs from approach to approach. Three classification models are used in our work. They are

1) K-Nearest Neighbor (K-NN)

2) Random Forest

3) Naïve Bayes

The following sections explain each of them briefly.

5.5.1 K-Nearest Neighbor Algorithm (K-NN)

This is one of the oldest, simplest and most popular classification algorithms used in data mining and pattern recognition field. K-NN algorithm [59] tries to find k-nearest points to the testing point and labels (maps) the testing point to a class which appeared maximum number of times in the circle enclosing k-training points. This algorithm is also called as Lazy learning method and it requires very less or no training. The main steps of the algorithms are

1) Define k-value

2) Train the set

3) Distance measurement

4) Labeling.

The Figures 23 and 24 shows a general approach to label (map, classify) a given data set (point) from a set of predefined labels.

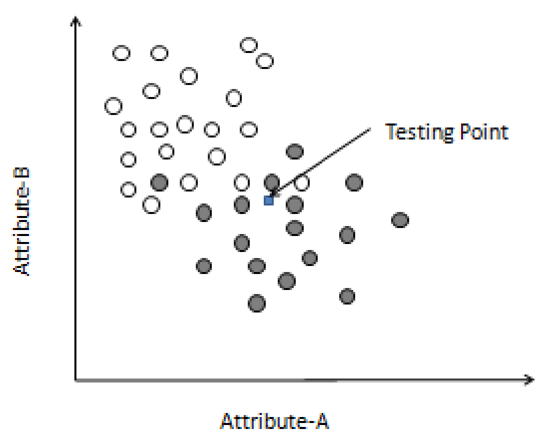

Figure 23 Figure showing scatter plot of two classes.

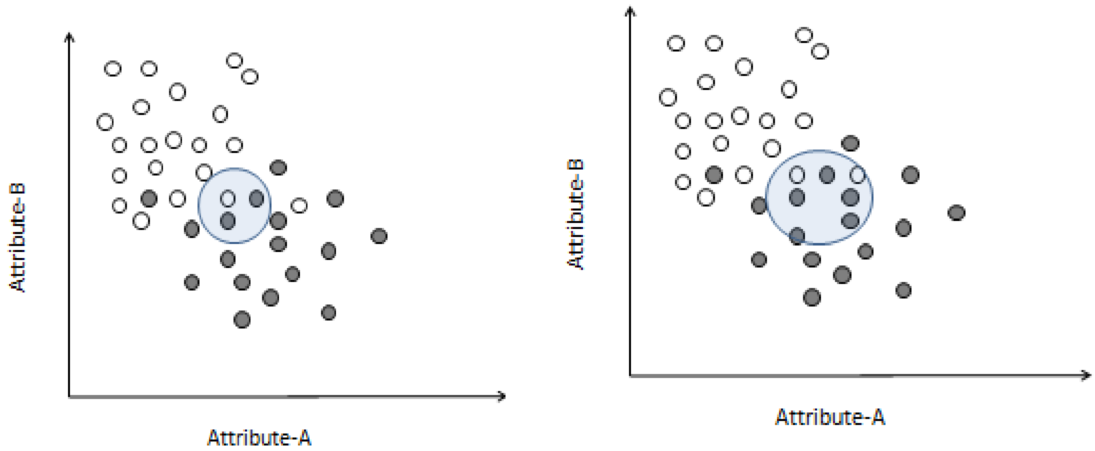

Figure 24 Figure showing decision making for different k-values.

Figure 23 shows a scatter-plot of two-classes, in which unfilled-'dots ' belong to Class YES and filled-'dots' belong to class NO. The square 'dot' is a point which is to be labeled. Actually the testing dot here belongs to the class YES.

Let's see how k-NN algorithm labels the testing set in the following section.

Case1. k=3: Here the value of k is 3, i.e. we are considering 3 nearest points to classify the testing set. We find that out of 3-nearest points, two dots belong to the class NO and one-point belongs to the class YES. Majority is the class NO. Hence the model labels the testing set as class NO.

Case2. k=7: Here the value of k is 7, i.e. we are considering 7 nearest points to classify the testing set. We find that out of 7-nearest points, five dots belong to the class NO and two-point belongs to the class YES. Majority is the class NO. Hence again the model labels the testing set as class NO.

As we have seen, more the number of k-values more the training set it uses for training the model. In the first case the probability of the testing set belonging to class NO was 0.66. But in the second case the probability of labeling the test set to class NO was increased to 0.71, which is more accurate. Hence we can understand a very important fact that, more the value of k more the accurate is the model.

5.5.2 Random Forest

Random forest is a type of recursive partitioning decision tree method which involve an ensemble of classification (or regression) trees that are calculated on random subsets of the data, using a subset of randomly restricted and selected predictors for each split in each classification tree In this way, random forests are able to better examine the contribution and behavior that each predictor has, even when one predictor’s effect would usually be overshadowed by more significant competitors in simpler models (e.g., simple or mixed effect regression models). Furthermore, the results of an ensemble of classification/regression trees have been shown to produce better predictions than the results of one classification tree on its own [60] [61].

5.5.3 Naïve Bayes

The Naive Bayes algorithm utilizes a very simple procedure to assign a class to a new instance. The probability of the existing classes is computed and the instance is put in the class having the highest probability which depends on the feature values present in the instance that has to be assigned a class.

The Naïve Bayes classifier technique is based on the Bayesian theorem and is particularly suited when the dimensionality of the inputs is high. Despite of its simplicity, Naïve Bayes can often outperform more sophisticated classification methods. To demonstrate the concept of Naïve Bayes classification, consider the example displayed in the Figure 25. The objects can be classified as either GREEN or RED. Our task is to classify new cases as they arrive, i.e., decide to which class label they belong, based on the currently existing objects.

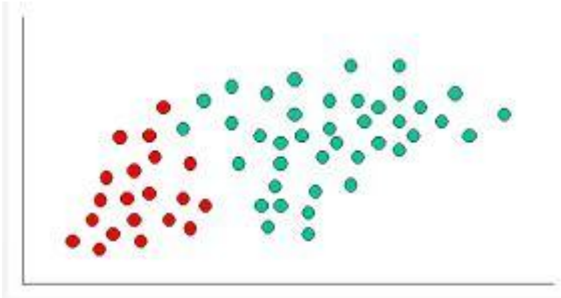

Figure 25 Naïve Bayes example illustration

Since there are twice as many GREEN objects as RED, it is reasonable to believe that a new case (which is unobserved yet) is twice as likely to have membership GREEN rather than RED. In the Bayesian analysis, this belief is known as the prior probability. Prior probabilities are based on previous experience, in this case the percentage of GREEN and RED objects, and often used to predict outcomes before they actually happen.

$$\text{Probability for the GREEN} = \frac{Number\ of\ Green\ objects}{Total\ number\ of\ objects}$$

$$\text{Probability for the RED} = \frac{Number\ of\ Red\ objects}{Total\ number\ of\ objects}$$

Since there are total of 60 objects, 40 of which are GREEN and 20 RED, our prior probabilities for class memberships are:

$$\text{Probability for the GREEN} = \frac{40}{60}$$

$$\text{Probability for the RED} = \frac{20}{60}$$

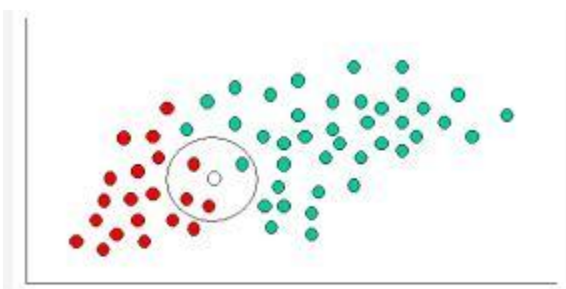

Figure 26 Probability calculated by Naïve Bayes

The naïve Bayes classifier can be adversely affected by redundant attributes due to its assumption that attributes are independent given the class. Its performance can be increased by removing the irrelevant and redundant features from the data set [62].

Chapter 6

Results

This chapter deals with the results of the work at various levels. The results are presented as:

1. Texture Feature Extraction
2. Dimensionality reduction
3. Classification
4. Classification into fatty and dense

6.1 FEATURE EXTRACTION

The texture features are measured for three classes – Fatty, Glandular and Dense. In all 13 GLCM features are extracted which are presented in the Table 2. The new proposed features which are extracted are shown in Table 3. The various ranges of the intensity which were taken during the experiments are also shown in the Table 3. The contradiction of the contrast between GLCM and proposed one is also shown in the table as the correlation between them (GLCM contrast and the normalized values of proposed contrast) was calculated which was 0.096. So both the contrasts were chosen for the data analysis work.

Table 2 Extracted GLCM Features.

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog- -eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb001	0.4968	0.0413	753.1339	44.0187	0.9793	13.2688	156.2091	0.8063	0.8348	0.1137	0.1722	-0.7397	0.7911
mdb002	0.8465	0.0106	459.1996	47.8033	0.9947	13.8513	183.2161	0.3258	0.3328	0.122	0.059	-0.8015	0.604
mdb003	0.8267	0.0264	461.1922	47.7272	0.9868	13.8396	181.8052	0.366	0.3843	0.1176	0.1222	-0.618	0.5424
mdb004	0.7706	0.0275	494.3947	47.3221	0.9863	13.7711	177.8419	0.4495	0.4683	0.1174	0.1258	-0.676	0.6205
mdb005	0.6431	0.1062	268.7504	23.4535	0.9469	9.69	82.0373	0.6556	0.7292	0.0979	0.3384	-0.3092	0.4839
mdb006	0.8291	0.0245	248.1643	25.6689	0.9878	10.1505	95.8643	0.3733	0.3903	0.1182	0.115	-0.6434	0.5581
mdb007	0.6665	0.0502	317.272	26.5353	0.9749	10.3063	93.7788	0.6503	0.685	0.1114	0.1991	-0.6081	0.6728
mdb008	0.5367	0.0716	407.8774	27.6957	0.9642	10.5163	94.7886	0.8187	0.8682	0.106	0.2577	-0.5827	0.715
mdb009	0.8515	0.0274	238.9989	25.5788	0.9863	10.1327	96.2588	0.3332	0.3522	0.1174	0.1257	-0.5691	0.4958
mdb010	0.3813	0.073	659.4145	29.4247	0.9635	10.8027	94.6835	1.1463	1.1959	0.1057	0.2613	-0.6874	0.8468
mdb011	0.4893	0.0987	504.4882	32.5292	0.9506	11.4025	111.7591	0.8664	0.9348	0.0996	0.3222	-0.4933	0.677
mdb012	0.8411	0.0349	238.1331	25.3736	0.9826	10.092	94.5451	0.38	0.4041	0.1154	0.1514	-0.5192	0.4991
mdb013	0.459	0.057	745.9402	41.1295	0.9715	12.817	143.5144	0.8762	0.9156	0.1096	0.2188	-0.6824	0.7835
mdb014	0.4776	0.0427	777.7082	43.5545	0.9786	13.1921	153.9646	0.821	0.8505	0.1133	0.1765	-0.7389	0.7946
mdb015	0.443	0.0856	557.8211	31.4532	0.9572	11.206	107.0323	0.9026	0.9619	0.1026	0.2922	-0.5679	0.7306
mdb016	0.6507	0.0686	283.7509	23.7524	0.9657	9.743	82.1793	0.7076	0.7551	0.1067	0.2501	-0.536	0.6529
mdb017	0.6961	0.0439	399.4686	34.7485	0.978	11.794	124.9605	0.6388	0.6691	0.113	0.1803	-0.6388	0.685
mdb018	0.9445	0.0171	301.8785	35.8975	0.9914	12.0115	140.3788	0.1659	0.1778	0.1202	0.0867	-0.3708	0.2813
mdb019	0.452	0.094	491.4643	28.8371	0.953	10.7269	97.418	0.8991	0.9642	0.1007	0.3116	-0.5306	0.7084
mdb020	0.754	0.0323	273.6522	25.6568	0.9839	10.137	92.515	0.5409	0.5629	0.1161	0.1425	-0.6832	0.6672
mdb021	0.8368	0.0339	338.3416	35.7857	0.9831	11.9859	134.2402	0.4102	0.4336	0.1157	0.1479	-0.5469	0.5298
mdb022	0.4505	0.0705	706.9551	38.5566	0.9648	12.395	129.9387	1.0467	1.0952	0.1063	0.2549	-0.6674	0.8176
mdb023	0.5595	0.0448	577.9332	39.2193	0.9776	12.5235	139.2167	0.7571	0.7878	0.1128	0.1829	-0.704	0.7595
mdb024	0.4943	0.0652	554.8208	33.5155	0.9674	11.5566	113.0286	0.9728	1.017	0.1076	0.241	-0.6637	0.8003
mdb025	0.4791	0.0696	313.1126	18.3944	0.9652	8.5513	58.7371	0.9502	0.9983	0.1065	0.2527	-0.6401	0.7815
mdb026	0.4489	0.0484	400.8128	20.8533	0.9758	9.0956	67.5986	0.9396	0.9726	0.1118	0.1938	-0.7352	0.824
mdb027	0.324	0.0611	481.6605	16.5415	0.9695	8.0461	48.0846	1.258	1.2996	0.1086	0.2299	-0.7626	0.8937

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog--eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb028	0.5538	0.0408	234.3413	15.4597	0.9796	7.8236	49.1292	0.8831	0.9103	0.1138	0.1705	-0.7557	0.8193
mdb029	0.6166	0.047	602.3496	45.9993	0.9765	13.5683	167.0074	0.6696	0.7021	0.1122	0.1895	-0.6532	0.704
mdb030	0.5902	0.0677	610.198	45.879	0.9662	13.5536	165.3787	0.718	0.7649	0.107	0.2475	-0.5584	0.6691
mdb031	0.449	0.0605	726.3566	38.6311	0.9697	12.4158	130.6534	1.0382	1.0798	0.1088	0.2284	-0.7096	0.8347
mdb032	0.5341	0.0621	559.1497	36.7106	0.9689	12.1027	125.4467	0.9452	0.9881	0.1083	0.2328	-0.6705	0.7944
mdb033	0.5133	0.0655	577.4558	36.0027	0.9672	11.9839	121.6989	1.0036	1.0488	0.1075	0.2418	-0.7014	0.8239
mdb034	0.4214	0.0421	615.2796	29.1078	0.9789	10.7502	96.0354	1.0176	1.0458	0.1135	0.1746	-0.7859	0.8618
mdb035	0.7021	0.0347	398.4227	34.4858	0.9826	11.7514	124.3029	0.6282	0.6521	0.1154	0.1508	-0.7048	0.7142
mdb036	0.6395	0.0422	442.0186	34.5647	0.9789	11.7487	121.3325	0.7687	0.796	0.1135	0.1749	-0.7073	0.7672
mdb037	0.4971	0.0359	758.9913	43.9555	0.982	13.252	156.2105	0.7897	0.8144	0.1151	0.1548	-0.7669	0.7986
mdb038	0.9998	0.0001	287.9001	35.8123	1	11.9999	143.9758	0.0009	0.001	0.125	0.0009	0	0.0001
mdb039	0.5213	0.0635	610.4149	39.4811	0.9683	12.564	138.421	0.8328	0.8768	0.108	0.2365	-0.6284	0.7438
mdb040	0.3851	0.0782	851.283	39	0.9609	12.4514	128.8764	1.1633	1.2157	0.1044	0.2744	-0.6814	0.848
mdb041	0.3882	0.0579	697.7203	30.3774	0.971	10.9905	98.2008	1.1612	1.2002	0.1094	0.2213	-0.7489	0.8732
mdb042	0.3932	0.0418	690.498	29.2239	0.9791	10.7569	94.4083	1.1238	1.1513	0.1136	0.1736	-0.8063	0.8891
mdb043	0.6417	0.0503	430.1398	34.3846	0.9749	11.7298	121.6979	0.7269	0.7617	0.1114	0.1994	-0.6452	0.7183
mdb044	0.8324	0.0293	345.8912	36.1667	0.9853	12.046	135.6006	0.4127	0.4329	0.1169	0.1323	-0.608	0.5643
mdb045	0.5019	0.0618	425.2959	25.873	0.9691	10.1488	84.922	0.9913	1.0341	0.1084	0.2319	-0.6858	0.8136
mdb046	0.4674	0.0593	335.0575	18.7919	0.9703	8.6382	60.208	0.944	0.985	0.1091	0.2251	-0.6865	0.8025
mdb047	0.4804	0.0557	452.5393	25.7692	0.9722	10.1242	84.0172	1.0225	1.061	0.11	0.2149	-0.7225	0.8366
mdb048	0.4712	0.0644	494.3428	28.2504	0.9678	10.6145	94.5635	0.9404	0.9849	0.1078	0.2389	-0.662	0.7896
mdb049	0.4725	0.0409	750.6939	41.3124	0.9796	12.8471	145.5059	0.8233	0.8517	0.1138	0.1708	-0.75	0.8001
mdb050	0.9063	0.0192	316.9964	36.0707	0.9904	12.0348	138.9052	0.2547	0.2679	0.1196	0.0949	-0.5651	0.4367
mdb051	0.6678	0.0472	406.9188	34.0069	0.9764	11.669	122.4994	0.6253	0.658	0.1122	0.1901	-0.6262	0.6718
mdb052	0.4411	0.06	579.069	30.5079	0.97	11.0275	103.2465	0.9151	0.9567	0.1089	0.2271	-0.679	0.7915
mdb053	0.427	0.0749	787.4009	41.3555	0.9626	12.8587	142.521	0.9623	1.0132	0.1052	0.266	-0.6321	0.7812
mdb054	0.73	0.0565	399.4392	37.4375	0.9717	12.2473	137.7597	0.5263	0.5653	0.1098	0.2173	-0.4848	0.5535
mdb055	0.5734	0.0478	552.7168	38.7321	0.9761	12.4427	136.7974	0.7792	0.8122	0.112	0.1921	-0.6857	0.7574
mdb056	0.4486	0.0598	531.7369	28.3695	0.9701	10.6226	93.7796	0.996	1.0369	0.1089	0.2264	-0.6987	0.8207

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog--eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb057	0.4472	0.065	573.7744	31.2101	0.9675	11.1569	106.1006	0.9021	0.9471	0.1076	0.2405	-0.6537	0.7756
mdb058	0.7093	0.0462	390.1549	34.8259	0.9769	11.8085	125.769	0.6153	0.6472	0.1124	0.1872	-0.618	0.6647
mdb059	0.5144	0.0848	274.6063	18.1291	0.9576	8.4981	58.7149	0.8869	0.9456	0.1028	0.2904	-0.5463	0.7145
mdb060	0.7027	0.051	177.7632	15.764	0.9745	7.9327	53.3573	0.6629	0.698	0.1112	0.2013	-0.601	0.6741
mdb061	0.781	0.0277	356.7433	34.5844	0.9862	11.7829	128.9401	0.444	0.4619	0.1173	0.1266	-0.6621	0.6137
mdb062	0.7187	0.0597	284.0689	26.2877	0.9702	10.2574	94.6903	0.5466	0.5876	0.109	0.2261	-0.4761	0.5578
mdb063	0.5364	0.0451	572.1729	36.4188	0.9774	12.0503	123.9708	0.9731	1.0041	0.1127	0.184	-0.7608	0.8426
mdb064	0.6098	0.0562	495.2344	37.6536	0.9719	12.267	132.1637	0.8027	0.8414	0.1098	0.2164	-0.647	0.744
mdb065	0.3162	0.0496	1207.613	37.8944	0.9752	12.2493	121.7121	1.3374	1.3683	0.1115	0.1974	-0.8119	0.921
mdb066	0.5007	0.0631	413.6815	25.1073	0.9685	9.9868	81.5927	1.0164	1.0595	0.1081	0.2354	-0.6947	0.8235
mdb067	0.7713	0.024	496.8865	47.1445	0.988	13.7481	176.5539	0.4793	0.4957	0.1183	0.1133	-0.7305	0.663
mdb068	0.5247	0.0457	711.5523	44.641	0.9772	13.3564	157.8587	0.8298	0.8615	0.1125	0.1856	-0.7247	0.7903
mdb069	0.4604	0.0521	369.6485	19.6671	0.974	8.8326	62.9149	0.979	1.0148	0.1109	0.2046	-0.7335	0.8321
mdb070	0.5078	0.0635	449.9723	28.3191	0.9682	10.633	97.2577	0.8118	0.8559	0.108	0.2366	-0.6287	0.7375
mdb071	0.4911	0.0435	653.8271	37.3569	0.9783	12.2034	127.139	0.9796	1.0097	0.1131	0.1789	-0.7697	0.8472
mdb072	0.384	0.0409	982.9276	40.4047	0.9796	12.6438	133.8511	1.1779	1.2046	0.1138	0.1708	-0.8205	0.9027
mdb073	0.438	0.0864	566.904	31.4387	0.9568	11.203	106.0768	0.9484	1.0081	0.1025	0.294	-0.5739	0.7464
mdb074	0.365	0.0689	717.7233	29.705	0.9655	10.8374	94.5264	1.2046	1.2511	0.1067	0.2509	-0.7198	0.8703
mdb075	0.8555	0.0292	236.0729	25.4504	0.9854	10.1065	95.4482	0.348	0.3683	0.1169	0.1319	-0.5417	0.4943
mdb076	0.6061	0.046	365.9942	27.2998	0.977	10.4523	96.1389	0.6807	0.7124	0.1125	0.1867	-0.664	0.7142
mdb077	0.7807	0.0324	245.1891	23.9275	0.9838	9.7977	87.7052	0.4511	0.4734	0.116	0.143	-0.6215	0.5925
mdb078	0.7109	0.0416	295.1123	26.2208	0.9792	10.2456	93.731	0.5888	0.6176	0.1136	0.173	-0.6331	0.6601
mdb079	0.481	0.0409	575.2735	32.1372	0.9796	11.3258	110.0368	0.8839	0.9121	0.1138	0.1706	-0.7602	0.8209
mdb080	0.4902	0.0634	471.6404	28.4367	0.9683	10.6525	96.7245	0.8616	0.9055	0.108	0.2363	-0.645	0.7606
mdb081	0.3654	0.0466	923.3335	36.0538	0.9767	11.9694	118.1851	1.1806	1.2125	0.1123	0.1884	-0.801	0.8957
mdb082	0.418	0.0699	566.7942	28.1902	0.965	10.5909	91.9378	1.0667	1.1151	0.1064	0.2535	-0.6795	0.8264
mdb083	0.8429	0.0266	335.4799	35.4597	0.9867	11.9308	133.5633	0.3845	0.4029	0.1176	0.1228	-0.6174	0.5523
mdb084	0.4276	0.0778	616.0252	32.4121	0.9611	11.3665	107.997	1.0273	1.0812	0.1045	0.2734	-0.6434	0.8013
mdb085	0.458	0.045	547.1575	28.9815	0.9775	10.7478	97.2492	0.9416	0.972	0.1127	0.1834	-0.7527	0.8326

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog--eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb086	0.477	0.0598	483.1709	27.7083	0.9701	10.5044	92.0681	0.9632	1.0044	0.1089	0.2265	-0.6867	0.8081
mdb087	0.4644	0.0753	494.8573	28.6649	0.9623	10.6888	96.3868	0.917	0.9691	0.1051	0.2671	-0.6078	0.7559
mdb088	0.5871	0.06	360.2194	26.352	0.97	10.2583	89.6048	0.8324	0.874	0.1089	0.2268	-0.6377	0.7476
mdb089	0.5456	0.0645	567.5497	38.6557	0.9677	12.4295	134.9055	0.8489	0.8935	0.1078	0.2392	-0.6269	0.7481
mdb090	0.6642	0.0477	558.8293	46.4773	0.9762	13.6474	170.4949	0.6105	0.6435	0.112	0.1917	-0.6155	0.6617
mdb091	0.4789	0.0573	419.7915	23.8873	0.9714	9.7502	77.5143	1.0112	1.0505	0.1096	0.2194	-0.7134	0.8303
mdb092	0.4185	0.0752	426.0595	21.7175	0.9624	9.28	69.1399	1.0329	1.0847	0.1051	0.2669	-0.6517	0.806
mdb093	0.4572	0.0595	460.3287	24.6768	0.9703	9.9039	78.9125	1.097	1.1374	0.109	0.2255	-0.7247	0.8529
mdb094	0.4667	0.0567	380.381	21.3605	0.9716	9.217	70.5972	0.8696	0.9086	0.1097	0.2179	-0.6793	0.7804
mdb095	0.5081	0.045	473.3187	28.621	0.9775	10.6834	98.1997	0.8193	0.8504	0.1127	0.1834	-0.7235	0.7873
mdb096	0.4369	0.0496	600.9867	30.1001	0.9752	10.9461	100.3866	0.9905	1.0248	0.1115	0.1974	-0.7481	0.8405
mdb097	0.7087	0.06	376.8732	34.337	0.97	11.7281	125.1809	0.558	0.5995	0.1089	0.2268	-0.489	0.5691
mdb098	0.3993	0.1292	535.1695	29.472	0.9354	10.8462	98.2926	0.9748	1.0643	0.0929	0.3848	-0.4376	0.6703
mdb099	0.6587	0.0486	418.8618	34.4565	0.9757	11.743	122.7057	0.6923	0.726	0.1118	0.1945	-0.6399	0.7041
mdb100	0.4737	0.08	542.8746	32.4265	0.96	11.3787	110.1053	0.928	0.9833	0.104	0.2789	-0.5915	0.7501
mdb101	0.4585	0.0546	598.3657	32.186	0.9727	11.3286	107.6148	1.0124	1.0499	0.1103	0.2118	-0.7265	0.8357
mdb102	0.342	0.0481	883.6391	30.8962	0.976	11.0147	95.9252	1.3537	1.3838	0.1119	0.1927	-0.8168	0.9241
mdb103	0.325	0.0542	794.8921	26.9085	0.9729	10.3148	84.3439	1.2451	1.2805	0.1104	0.2107	-0.7864	0.9005
mdb104	0.451	0.0503	617.9309	31.8058	0.9749	11.226	104.5493	1.0782	1.1108	0.1114	0.1994	-0.7635	0.8654
mdb105	0.4651	0.0374	1057.473	56.5967	0.9813	15.0485	203.2464	0.8257	0.8514	0.1147	0.1598	-0.7693	0.8095
mdb106	0.5015	0.0303	768.0389	43.9499	0.9849	13.2419	154.5883	0.8527	0.8731	0.1166	0.1358	-0.8089	0.8352
mdb107	0.7669	0.0267	497.3111	47.2612	0.9867	13.7653	177.5824	0.4538	0.4723	0.1176	0.123	-0.6934	0.6282
mdb108	0.4545	0.03	837.2252	42.4817	0.985	13.0186	147.9757	0.9016	0.9223	0.1167	0.1348	-0.8216	0.8511
mdb109	0.7588	0.0362	380.8508	36.1123	0.9819	12.0358	132.1784	0.5571	0.5818	0.115	0.1555	-0.6513	0.6561
mdb110	0.572	0.0335	484.1734	32.949	0.9832	11.4712	116.1801	0.7285	0.7513	0.1157	0.1467	-0.7566	0.7757
mdb111	0.587	0.0522	458.5503	33.3174	0.9739	11.5425	117.4476	0.7371	0.7732	0.1109	0.205	-0.6485	0.7238
mdb112	0.4823	0.0532	498.9547	28.9047	0.9734	10.7432	99.112	0.8322	0.869	0.1106	0.2078	-0.6872	0.7733
mdb113	0.4887	0.0587	536.827	31.961	0.9706	11.2946	110.2192	0.8372	0.8779	0.1092	0.2234	-0.6607	0.7615
mdb114	0.6475	0.0296	428.9316	33.5744	0.9852	11.5865	120.6796	0.6311	0.6513	0.1168	0.1333	-0.7513	0.7394

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog- -eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb115	0.4957	0.0477	556.2109	32.4628	0.9762	11.3736	110.3482	0.9177	0.9503	0.112	0.1915	-0.7316	0.8168
mdb116	0.2936	0.0453	1306.908	36.3309	0.9773	11.9771	115.0808	1.386	1.4148	0.1126	0.1845	-0.8329	0.9319
mdb117	0.4385	0.0664	585.0972	31.028	0.9668	11.1255	104.49	0.9531	0.9991	0.1073	0.2442	-0.6624	0.7928
mdb118	0.4021	0.0499	715.7263	31.4561	0.975	11.1734	102.1759	1.1441	1.1787	0.1114	0.1983	-0.783	0.8836
mdb119	0.6538	0.055	449.4519	36.9612	0.9725	12.1583	131.0758	0.7362	0.7741	0.1102	0.213	-0.6188	0.708
mdb120	0.5865	0.0512	512.092	36.756	0.9744	12.1203	127.4829	0.8689	0.9038	0.1111	0.202	-0.6968	0.7887
mdb121	0.6036	0.0431	523.0314	38.5813	0.9785	12.4267	137.7908	0.7177	0.7475	0.1132	0.1776	-0.6899	0.7396
mdb122	0.4155	0.0634	725.0671	35.0518	0.9683	11.8104	115.4818	1.1303	1.1743	0.108	0.2362	-0.7208	0.8568
mdb123	0.748	0.0481	362.9814	34.5449	0.976	11.7671	127.3486	0.4984	0.5317	0.1119	0.1929	-0.5262	0.5635
mdb124	0.4966	0.0349	761.072	43.865	0.9826	13.2462	156.0959	0.7886	0.8125	0.1154	0.1514	-0.7723	0.8008
mdb125	0.4001	0.0569	843.5971	38.0679	0.9716	12.3093	125.8477	1.1641	1.203	0.1097	0.2183	-0.7526	0.876
mdb126	0.6167	0.0501	476.0374	36.2394	0.9749	12.0395	126.9063	0.8075	0.8421	0.1114	0.1988	-0.6781	0.7617
mdb127	0.4664	0.0471	567.9261	31.1695	0.9764	11.155	107.2356	0.8448	0.8774	0.1122	0.1899	-0.7222	0.7933
mdb128	0.3827	0.0564	671.4989	28.5694	0.9718	10.6366	92.0192	1.1253	1.1636	0.1098	0.2169	-0.752	0.8693
mdb129	0.7148	0.0397	420.265	37.5241	0.9802	12.2616	137.2704	0.5653	0.5926	0.1141	0.1669	-0.6406	0.656
mdb130	0.4583	0.0488	794.4651	42.9279	0.9756	13.0969	150.9313	0.8499	0.8837	0.1117	0.1949	-0.7172	0.7923
mdb131	0.9423	0.0219	207.1514	24.6911	0.9891	9.9653	96.3333	0.153	0.1681	0.1189	0.1053	-0.2237	0.2044
mdb132	0.9954	0.0018	200.6251	24.8592	0.9991	10.0028	99.7125	0.0173	0.0186	0.1245	0.0131	-0.2318	0.0697
mdb133	0.9843	0.0039	203.1046	24.9088	0.998	10.0118	99.2284	0.0514	0.0542	0.1239	0.0257	-0.5033	0.1892
mdb134	0.5854	0.0861	298.9194	22.9263	0.957	9.5731	78.717	0.7339	0.7935	0.1025	0.2934	-0.4707	0.622
mdb135	0.9341	0.0216	301.182	35.5674	0.9892	11.9542	138.95	0.1694	0.1843	0.119	0.1041	-0.3059	0.2564
mdb136	0.5188	0.103	400.7479	27.7292	0.9485	10.5283	94.8681	0.8229	0.8943	0.0986	0.3317	-0.4501	0.6371
mdb137	0.4583	0.0496	722.8799	38.6278	0.9752	12.4093	131.2444	1.0056	1.04	0.1115	0.1974	-0.7509	0.8451
mdb138	0.378	0.0366	1039.372	40.8609	0.9817	12.7379	136.6169	1.1407	1.1653	0.1149	0.1568	-0.8347	0.9015
mdb139	0.4606	0.0918	481.2895	28.6969	0.9541	10.7025	97.1286	0.8889	0.9525	0.1012	0.3066	-0.5346	0.7077
mdb140	0.914	0.0199	219.5046	25.2204	0.99	10.0689	97.2674	0.2122	0.226	0.1194	0.0977	-0.4905	0.3701
mdb141	0.9462	0.0103	212.3851	25.0867	0.9949	10.0444	98.0803	0.1447	0.1518	0.1221	0.0573	-0.5834	0.344
mdb142	0.807	0.0322	254.6537	25.8216	0.9839	10.1776	95.8863	0.4003	0.4226	0.1161	0.1422	-0.5885	0.5458
mdb143	0.4351	0.0915	523.1649	29.3061	0.9542	10.8091	98.7542	0.9156	0.9789	0.1012	0.3061	-0.5486	0.7236

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog- -eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb144	0.5875	0.0533	376.1682	27.4328	0.9733	10.4746	96.0183	0.71	0.7469	0.1106	0.2082	-0.6351	0.7088
mdb145	0.3284	0.0659	1012.214	35.9494	0.967	11.9507	116.2356	1.2595	1.3042	0.1074	0.2429	-0.7452	0.8875
mdb146	0.4648	0.0485	707.0019	38.2774	0.9757	12.3526	129.9468	1.006	1.0396	0.1118	0.1942	-0.7529	0.8459
mdb147	0.724	0.0461	287.6807	26.2444	0.9769	10.2519	94.0991	0.5755	0.6075	0.1124	0.187	-0.5923	0.6327
mdb148	0.4416	0.0582	565.4255	29.5288	0.9709	10.8543	98.9666	0.9591	0.9991	0.1093	0.222	-0.6969	0.8105
mdb149	0.4578	0.061	433.1623	23.6257	0.9695	9.6818	75.9233	1.037	1.0789	0.1086	0.2297	-0.7063	0.8327
mdb150	0.4934	0.053	367.4182	21.7887	0.9735	9.3191	72.715	0.8437	0.8803	0.1107	0.2073	-0.6872	0.7764
mdb151	0.6783	0.0643	433.2077	37.8274	0.9679	12.3094	137.5262	0.6019	0.6464	0.1078	0.2385	-0.4984	0.5914
mdb152	0.4173	0.1146	776.6091	43.5255	0.9427	13.1937	150.796	0.9484	1.0277	0.096	0.356	-0.4752	0.6882
mdb153	0.9283	0.021	135.5415	15.8371	0.9896	7.9855	60.6108	0.2058	0.2206	0.1192	0.1013	-0.4167	0.3342
mdb154	0.6064	0.0378	309.2136	22.7771	0.9811	9.5345	79.2331	0.6713	0.6973	0.1146	0.1609	-0.7125	0.7352
mdb155	0.5841	0.0791	441.229	33.3903	0.9605	11.5601	117.8554	0.732	0.7868	0.1042	0.2765	-0.5054	0.6431
mdb156	0.4232	0.0876	553.5702	29.7303	0.9562	10.8859	99.681	0.9484	1.0089	0.1022	0.297	-0.5736	0.7461
mdb157	0.5786	0.0823	362.3207	27.2615	0.9588	10.4447	94.7841	0.7404	0.7973	0.1034	0.2844	-0.4929	0.639
mdb158	0.8213	0.039	246.1554	25.6832	0.9805	10.1528	95.735	0.3809	0.4079	0.1143	0.1648	-0.4885	0.4817
mdb159	0.797	0.0375	348.5254	35.0912	0.9812	11.8628	130.3948	0.4572	0.4832	0.1147	0.16	-0.5715	0.5681
mdb160	0.4423	0.072	541.4336	29.5415	0.964	10.8593	100.0012	0.9056	0.9554	0.1059	0.2588	-0.6267	0.7631
mdb161	0.5906	0.0455	514.0948	36.9037	0.9773	12.1479	128.7945	0.8358	0.8668	0.1126	0.1849	-0.7127	0.788
mdb162	0.4078	0.0631	866.3936	41.5569	0.9684	12.871	141.0624	1.0466	1.0901	0.1081	0.2355	-0.7048	0.8343
mdb163	0.4597	0.0514	759.0802	41.341	0.9743	12.8554	144.7358	0.864	0.8995	0.1111	0.2026	-0.7057	0.7908
mdb164	0.6559	0.0583	557.7499	46.5197	0.9708	13.6493	170.0516	0.6289	0.6693	0.1093	0.2224	-0.5566	0.6361
mdb165	0.7486	0.0236	377.0275	34.6226	0.9882	11.7732	126.8679	0.5316	0.5469	0.1184	0.1118	-0.7535	0.6985
mdb166	0.7078	0.0378	403.8534	35.4027	0.9811	11.9117	127.6144	0.6388	0.6644	0.1146	0.1608	-0.6904	0.7118
mdb167	0.423	0.1047	351.0268	19.545	0.9477	8.8175	62.8116	0.9473	1.0198	0.0982	0.3352	-0.5094	0.7085
mdb168	0.471	0.0827	362.4898	21.7662	0.9586	9.3095	71.9699	0.8748	0.9321	0.1033	0.2854	-0.5646	0.7215
mdb169	0.5638	0.0583	486.7077	33.9665	0.9709	11.6474	117.5197	0.8437	0.8839	0.1093	0.2221	-0.6532	0.7592
mdb170	0.4305	0.0502	608.9825	29.7747	0.9749	10.8897	98.6917	1.0169	1.0514	0.1114	0.199	-0.7492	0.8463
mdb171	0.9929	0.0028	394.0976	48.8143	0.9986	14.0044	195.4112	0.0256	0.0276	0.1242	0.0191	-0.2345	0.0854
mdb172	0.9744	0.0029	401.1099	48.6401	0.9986	13.9774	193.2744	0.0766	0.0784	0.1242	0.0196	-0.7398	0.303

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog- -eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb173	0.4753	0.0636	495.2284	28.8448	0.9682	10.7358	98.634	0.8481	0.8919	0.108	0.2368	-0.6425	0.756
mdb174	0.6036	0.043	370.4234	27.4078	0.9785	10.462	96.3491	0.6803	0.7097	0.1132	0.1774	-0.6826	0.7239
mdb175	0.7932	0.0376	348.4246	34.8757	0.9812	11.8303	129.87	0.4479	0.4738	0.1147	0.1602	-0.5599	0.5568
mdb176	0.8382	0.0267	336.8573	35.407	0.9866	11.9154	133.1175	0.3889	0.4064	0.1176	0.1231	-0.6213	0.5611
mdb177	0.5607	0.0456	564.9486	38.266	0.9772	12.364	133.7154	0.8374	0.8686	0.1126	0.1853	-0.714	0.7892
mdb178	0.6479	0.0483	456.9057	36.7578	0.9758	12.1269	130.174	0.7463	0.7798	0.1119	0.1935	-0.6647	0.7363
mdb179	0.5078	0.0698	837.7635	53.4108	0.9651	14.6185	191.2582	0.8173	0.8657	0.1064	0.2531	-0.6004	0.724
mdb180	0.9534	0.0104	408.7881	48.7782	0.9948	13.9972	191.9834	0.1437	0.1509	0.1221	0.0579	-0.5525	0.3313
mdb181	0.4846	0.0455	506.1017	28.9567	0.9772	10.7507	99.3016	0.8314	0.8629	0.1126	0.1851	-0.7247	0.7908
mdb182	0.531	0.0493	436.0458	28.0125	0.9754	10.5709	96.1395	0.8079	0.8418	0.1116	0.1963	-0.6918	0.7684
mdb183	0.704	0.0418	301.0317	26.4871	0.9791	10.3014	95.4885	0.5537	0.5826	0.1136	0.1736	-0.6234	0.6416
mdb184	0.5216	0.0809	416.7306	27.929	0.9596	10.562	95.844	0.8092	0.8653	0.1038	0.2808	-0.5429	0.6895
mdb185	0.4306	0.0911	525.3333	28.4993	0.9546	10.6611	93.9267	1.022	1.0859	0.1014	0.3053	-0.5812	0.7682
mdb186	0.4424	0.0563	599.8078	31.0627	0.9718	11.1204	103.8797	0.9846	1.0235	0.1098	0.2168	-0.7129	0.8239
mdb187	0.6632	0.0393	422.5341	34.472	0.9804	11.7499	123.0894	0.683	0.7099	0.1142	0.1657	-0.6934	0.7297
mdb188	0.4887	0.0505	609.9953	34.9219	0.9747	11.7915	117.4494	1.0092	1.0434	0.1113	0.2	-0.7506	0.8461
mdb189	0.5099	0.0544	263.448	16.1073	0.9729	7.9809	50.1474	0.9759	1.0135	0.1103	0.2113	-0.719	0.8247
mdb190	0.4041	0.1187	235.1349	12.8521	0.9407	7.1321	38.8852	0.966	1.0482	0.0951	0.3642	-0.4695	0.6887
mdb191	0.6418	0.0591	569.1148	46.3816	0.9704	13.6284	169.0458	0.6476	0.6886	0.1091	0.2246	-0.5644	0.6474
mdb192	0.8497	0.0288	334.2721	35.7584	0.9856	11.9797	134.7447	0.3817	0.4016	0.117	0.1304	-0.5795	0.5308
mdb193	0.662	0.0212	582.3346	46.2289	0.9894	13.6047	170.2494	0.5802	0.5948	0.1191	0.1027	-0.8016	0.7418
mdb194	0.4954	0.0468	747.1211	44.1377	0.9766	13.2831	156.5228	0.8071	0.8396	0.1122	0.1891	-0.7131	0.7783
mdb195	0.4638	0.0522	570.0302	31.4005	0.9739	11.1886	107.377	0.8719	0.9078	0.1109	0.205	-0.702	0.7913
mdb196	0.5596	0.0404	420.3048	28.0002	0.9798	10.5728	96.3001	0.8077	0.8354	0.1139	0.1692	-0.74	0.7922
mdb197	0.5834	0.0435	641.5045	45.5298	0.9788	13.4969	164.3876	0.7033	0.7358	0.1134	0.1773	-0.6993	0.7399
mdb198	0.9152	0.0187	310.0453	35.6415	0.9907	11.9655	137.7459	0.234	0.2469	0.1198	0.0928	-0.5345	0.4065
mdb199	0.7639	0.0379	376.6489	36.0954	0.981	12.031	132.2224	0.5495	0.5757	0.1146	0.1613	-0.6263	0.6418
mdb200	0.4335	0.054	680.3331	33.7246	0.973	11.5737	111.1613	1.0967	1.1317	0.1104	0.21	-0.7508	0.8647
mdb201	0.7077	0.0465	393.4652	34.9956	0.9767	11.8419	126.1704	0.6313	0.6635	0.1123	0.1882	-0.6202	0.672

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog- -eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb202	0.4883	0.0563	613.602	35.5185	0.9719	11.8912	119.2044	1.0286	1.0661	0.1098	0.2166	-0.7325	0.8434
mdb203	0.4329	0.0725	566.0856	30.1954	0.9637	10.9807	102.4914	0.9028	0.9529	0.1058	0.2601	-0.6248	0.7612
mdb204	0.4544	0.0744	496.072	27.8232	0.9628	10.5221	91.7237	0.9998	1.0503	0.1053	0.2648	-0.649	0.7988
mdb205	0.7168	0.048	260.5852	23.6059	0.976	9.726	84.7302	0.5444	0.5776	0.112	0.1925	-0.565	0.6062
mdb206	0.5786	0.0792	306.4047	22.9302	0.9604	9.5691	78.133	0.7655	0.8204	0.1042	0.2769	-0.5134	0.6585
mdb207	0.4497	0.0634	749.1599	40.7916	0.9683	12.7629	141.0338	0.9285	0.9725	0.108	0.2362	-0.6708	0.7909
mdb208	0.4454	0.0524	809.1621	42.4464	0.9738	13.0196	147.7782	0.9047	0.9407	0.1108	0.2056	-0.7103	0.8046
mdb209	0.3567	0.0733	475.1227	19.327	0.9634	8.7307	58.3434	1.2074	1.2568	0.1056	0.2621	-0.7076	0.866
mdb210	0.4274	0.067	405.7073	20.8036	0.9665	9.0786	66.3606	1.0044	1.0503	0.1071	0.2458	-0.6772	0.8131
mdb211	0.4486	0.06	796.4966	43.0065	0.97	13.1046	150.1998	0.8876	0.929	0.1089	0.2268	-0.6719	0.7812
mdb212	0.4178	0.0447	883.1431	41.5421	0.9776	12.8583	141.5669	1.0159	1.0444	0.1128	0.1827	-0.7712	0.8573
mdb213	0.4236	0.0821	538.9845	28.0455	0.9589	10.5644	91.3648	1.0661	1.123	0.1035	0.2839	-0.6306	0.8032
mdb214	0.3893	0.0712	643.5509	29.4661	0.9644	10.8153	95.6669	1.1053	1.1534	0.1061	0.2567	-0.6861	0.8391
mdb215	0.4818	0.0474	712.0722	40.6453	0.9763	12.7447	142.5303	0.8445	0.8772	0.1121	0.1907	-0.7163	0.7914
mdb216	0.7545	0.0265	505.7195	47.1696	0.9868	13.7488	176.6996	0.4715	0.4897	0.1176	0.1223	-0.7017	0.6437
mdb217	0.46	0.0564	738.7935	40.4725	0.9718	12.7073	139.6984	0.9307	0.9697	0.1098	0.217	-0.6981	0.8043
mdb218	0.3928	0.0484	965.1326	41.4682	0.9758	12.8456	139.5834	1.1027	1.1362	0.1118	0.1938	-0.7795	0.8755
mdb219	0.6544	0.0381	421.0304	33.8918	0.9809	11.6444	121.6643	0.6412	0.6674	0.1145	0.1619	-0.6923	0.7136
mdb220	0.4267	0.0655	599.4002	30.4752	0.9672	11.02	101.7886	0.9847	1.0302	0.1075	0.242	-0.6732	0.8055
mdb221	0.4059	0.0631	859.7104	40.9277	0.9684	12.7752	138.6589	1.054	1.0977	0.1081	0.2355	-0.708	0.837
mdb222	0.4226	0.0545	830.4803	40.7479	0.9728	12.75	139.0313	1.0103	1.0477	0.1103	0.2114	-0.7305	0.838
mdb223	0.5148	0.0409	527.0567	32.1145	0.9796	11.3276	112.0012	0.7851	0.8134	0.1138	0.1706	-0.7363	0.783
mdb224	0.4235	0.0536	635.2901	30.778	0.9732	11.0433	100.6532	1.0921	1.1283	0.1105	0.2091	-0.7523	0.8637
mdb225	0.527	0.0451	539.5393	33.5349	0.9775	11.5616	113.973	0.9369	0.968	0.1127	0.1837	-0.7499	0.8295
mdb226	0.8031	0.0334	352.4737	35.5778	0.9833	11.944	131.9402	0.4714	0.4945	0.1158	0.1464	-0.6114	0.5965
mdb227	0.8338	0.0289	337.3594	35.3241	0.9855	11.9074	132.7176	0.3983	0.4183	0.117	0.1309	-0.5988	0.5505
mdb228	0.5869	0.043	466.3256	33.0927	0.9785	11.4894	115.6932	0.7734	0.8023	0.1132	0.1775	-0.7109	0.7699
mdb229	0.4189	0.0609	618.3682	30.5426	0.9696	11.0293	101.7053	1.0035	1.0457	0.1087	0.2294	-0.7065	0.8252
mdb230	0.4836	0.0418	561.8914	31.7787	0.9791	11.2591	109.3665	0.8464	0.8752	0.1136	0.1736	-0.7464	0.8049

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog- -eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb231	0.9139	0.0229	140.3723	16.1686	0.9885	8.0659	61.7817	0.2124	0.2283	0.1186	0.1092	-0.4213	0.3397
mdb232	0.4699	0.0791	320.4558	19.051	0.9604	8.7073	62.2959	0.8676	0.9225	0.1042	0.2767	-0.5798	0.7276
mdb233	0.4629	0.0449	588.9828	31.6788	0.9776	11.2388	107.8242	0.9044	0.9355	0.1128	0.1831	-0.7454	0.8192
mdb234	0.2987	0.0523	927.451	26.468	0.9739	10.1149	79.5287	1.4122	1.4445	0.1108	0.2051	-0.8176	0.9309
mdb235	0.4326	0.0522	617.5876	30.747	0.9739	11.0725	102.519	1.007	1.0425	0.1109	0.205	-0.7374	0.8396
mdb236	0.4469	0.0562	794.5095	42.4194	0.9719	13.0132	148.3163	0.8736	0.9123	0.1098	0.2164	-0.6877	0.7849
mdb237	0.5643	0.0472	403.1913	27.7598	0.9764	10.5324	96.7942	0.7314	0.7641	0.1122	0.1901	-0.682	0.7395
mdb238	0.5214	0.0678	429.1465	28.0527	0.9661	10.5815	96.2129	0.812	0.8588	0.1069	0.2479	-0.6028	0.7237
mdb239	0.7937	0.0234	482.2966	47.4266	0.9883	13.7956	179.4565	0.4123	0.428	0.1185	0.1108	-0.6919	0.6089
mdb240	0.9955	0.0013	393.1296	48.7658	0.9994	13.9968	195.4291	0.0174	0.0181	0.1246	0.0099	-0.3855	0.1045
mdb241	0.4613	0.0701	713.9442	40.8331	0.9649	12.7697	142.488	0.8697	0.9181	0.1064	0.2539	-0.6212	0.7508
mdb242	0.605	0.0528	481.2015	35.9651	0.9736	11.9907	125.4929	0.823	0.8596	0.1107	0.2065	-0.6833	0.7689
mdb243	0.9309	0.0111	416.3696	48.3829	0.9944	13.9401	189.5677	0.1755	0.1832	0.1219	0.0612	-0.6263	0.3975
mdb244	0.6559	0.0447	568.1788	46.3821	0.9776	13.6302	169.8662	0.6184	0.6494	0.1128	0.1827	-0.6431	0.6782
mdb245	0.5685	0.0562	391.1836	27.6233	0.9719	10.5069	96.1749	0.7358	0.7747	0.1099	0.2162	-0.6326	0.7154
mdb246	0.3027	0.0694	877.6816	27.6757	0.9653	10.3808	83.9034	1.3993	1.4435	0.1066	0.252	-0.7708	0.9158
mdb247	0.5926	0.0664	362.1263	27.2803	0.9668	10.4439	95.2777	0.7149	0.7609	0.1073	0.2441	-0.5638	0.6708
mdb248	0.6922	0.0593	298.6382	26.4552	0.9703	10.2957	94.8257	0.5807	0.6218	0.1091	0.2251	-0.5138	0.5924
mdb249	0.6705	0.0504	411.4926	34.6573	0.9748	11.7762	123.6853	0.68	0.7149	0.1113	0.1996	-0.6266	0.6935
mdb250	0.5881	0.0576	467.3842	34.2095	0.9712	11.6954	119.1833	0.8124	0.8523	0.1095	0.2204	-0.6452	0.7459
mdb251	0.5005	0.0682	352.6512	21.9777	0.9659	9.3652	73.5561	0.8361	0.8832	0.1068	0.2489	-0.611	0.7359
mdb252	0.6045	0.0608	356.7914	27.2167	0.9697	10.4323	95.3863	0.6975	0.7399	0.1087	0.2292	-0.5858	0.677
mdb253	0.9449	0.0143	408.7961	48.5099	0.9928	13.9581	190.7998	0.1475	0.1575	0.121	0.075	-0.4513	0.2961
mdb254	0.4528	0.0753	771.7734	43.6973	0.9623	13.2121	152.8702	0.8836	0.9358	0.1051	0.2673	-0.6036	0.7452
mdb255	0.4786	0.1045	443.6568	28.2568	0.9477	10.62	95.7681	0.8724	0.9448	0.0983	0.3349	-0.4733	0.6663
mdb256	0.3554	0.1754	567.8655	30.3583	0.9123	11.0047	100.2007	1.0359	1.1575	0.0837	0.4643	-0.3301	0.606
mdb257	0.7265	0.0202	531.0835	46.9362	0.9899	13.7134	174.7225	0.5138	0.5277	0.1194	0.0987	-0.7836	0.7026
mdb258	0.9022	0.0201	432.1162	48.9716	0.99	14.0235	189.4267	0.2654	0.2791	0.1194	0.0984	-0.5682	0.448
mdb259	0.5971	0.0443	714.1069	52.2927	0.9779	14.4658	190.2029	0.6991	0.7297	0.1129	0.1814	-0.6824	0.7288

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog--eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb260	0.6073	0.0545	684.8732	52.0282	0.9728	14.4339	189.4493	0.6928	0.7305	0.1103	0.2115	-0.6177	0.6925
mdb261	0.4813	0.0386	791.0495	43.6875	0.9807	13.1969	152.3152	0.902	0.9276	0.1144	0.1635	-0.7757	0.834
mdb262	0.6021	0.0343	628.9926	45.6284	0.9828	13.52	165.7913	0.6711	0.6943	0.1155	0.1494	-0.7352	0.7467
mdb263	0.431	0.0671	515.6351	26.3264	0.9665	10.2267	84.6225	1.099	1.1455	0.1071	0.246	-0.6981	0.8413
mdb264	0.3495	0.0651	457.1697	17.5533	0.9674	8.2833	50.8938	1.2856	1.3297	0.1076	0.2409	-0.7471	0.8926
mdb265	0.6586	0.045	461.1285	37.7775	0.9775	12.2947	135.7114	0.6704	0.7014	0.1127	0.1834	-0.6586	0.7074
mdb266	0.6056	0.0482	446.9446	33.4286	0.9759	11.5651	118.6097	0.7046	0.738	0.1119	0.1931	-0.657	0.7177
mdb267	0.6893	0.0474	406.13	35.0604	0.9763	11.8485	125.5247	0.6689	0.7017	0.1121	0.1907	-0.636	0.6933
mdb268	0.6879	0.0409	442.7973	37.9156	0.9796	12.3261	138.6192	0.5739	0.6022	0.1138	0.1706	-0.6437	0.661
mdb269	0.7941	0.0391	355.4926	35.6955	0.9804	11.9616	131.8421	0.4935	0.5204	0.1143	0.1651	-0.5769	0.5878
mdb270	0.485	0.0666	538.7491	32.1929	0.9667	11.3414	110.3131	0.88	0.9261	0.1072	0.2448	-0.6317	0.7588
mdb271	0.4665	0.129	342.494	22.1325	0.9355	9.3972	73.0137	0.8942	0.9836	0.0929	0.3845	-0.3924	0.6179
mdb272	0.4482	0.0921	371.1699	21.5081	0.9539	9.2525	70.589	0.9011	0.9649	0.1011	0.3074	-0.5393	0.7143
mdb273	0.8208	0.0486	233.5433	24.6118	0.9757	9.9399	90.6251	0.4323	0.466	0.1118	0.1943	-0.4375	0.4799
mdb274	0.4589	0.0514	369.6885	19.7725	0.9743	8.8561	63.8918	0.9319	0.9674	0.1111	0.2025	-0.7256	0.8175
mdb275	0.5344	0.0539	512.2199	33.3344	0.973	11.5421	115.0118	0.858	0.8952	0.1104	0.2099	-0.6836	0.7785
mdb276	0.4527	0.0645	565.4068	31.198	0.9678	11.1611	106.7422	0.8736	0.9183	0.1078	0.2391	-0.6494	0.7661
mdb277	0.4756	0.0593	550.9255	31.7293	0.9704	11.2545	108.6435	0.8753	0.9163	0.1091	0.225	-0.6678	0.7761
mdb278	0.4307	0.0642	594.5698	30.5449	0.9679	11.0295	102.1737	0.9746	1.019	0.1078	0.2384	-0.6755	0.8046
mdb279	0.6023	0.0447	524.4189	38.8318	0.9776	12.4614	139.3782	0.6837	0.7146	0.1128	0.1827	-0.6738	0.7202
mdb280	0.6219	0.0264	521.0699	38.6405	0.9868	12.4409	139.5719	0.656	0.6735	0.1177	0.1218	-0.7807	0.7628
mdb281	0.4375	0.0638	793.4538	41.9998	0.9681	12.9469	145.78	0.9131	0.9569	0.1079	0.2374	-0.6641	0.7842
mdb282	0.6094	0.0592	485.5008	37.0229	0.9704	12.1768	130.0016	0.8064	0.8468	0.1091	0.2249	-0.6391	0.7408
mdb283	0.645	0.0399	469.5721	37.0945	0.98	12.1841	131.808	0.733	0.7606	0.1141	0.1677	-0.7114	0.7552
mdb284	0.6745	0.0312	422.1212	34.6751	0.9845	11.7826	124.2586	0.6632	0.6854	0.1164	0.1387	-0.7449	0.7474
mdb285	0.7106	0.0357	392.3421	34.4754	0.9822	11.754	124.9542	0.5988	0.6229	0.1152	0.1539	-0.6796	0.6913
mdb286	0.514	0.0607	444.5058	28.1117	0.9697	10.5889	95.7839	0.8445	0.8864	0.1087	0.2288	-0.6461	0.756
mdb287	0.3675	0.068	953.8605	40.1076	0.966	12.6408	133.3436	1.1618	1.2081	0.1069	0.2483	-0.7172	0.8618
mdb288	0.7298	0.02	528.8841	46.9067	0.99	13.703	174.3004	0.5205	0.5332	0.1194	0.098	-0.78	0.7092

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog- -eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb289	0.4654	0.0529	712.7283	39.0944	0.9735	12.4802	133.2386	0.9863	1.0221	0.1107	0.2071	-0.7264	0.8309
mdb290	0.4997	0.06	638.8356	38.8727	0.97	12.4633	134.1194	0.9238	0.9646	0.1089	0.227	-0.681	0.7961
mdb291	0.6512	0.0328	582.1739	46.2406	0.9836	13.6047	169.4261	0.6115	0.6341	0.1159	0.1445	-0.7219	0.717
mdb292	0.4038	0.0486	900.0591	40.7319	0.9757	12.735	137.6858	1.0612	1.0947	0.1118	0.1945	-0.7654	0.8626
mdb293	0.5716	0.0687	362.3064	26.0043	0.9657	10.1823	87.1666	0.8897	0.9371	0.1067	0.2502	-0.6171	0.7538
mdb294	0.9069	0.0148	223.4128	25.2821	0.9926	10.0803	97.2393	0.2265	0.2367	0.1208	0.0769	-0.63	0.4437
mdb295	0.7722	0.0318	490.7199	47.3448	0.9841	13.7784	177.9737	0.4513	0.4734	0.1162	0.1408	-0.6404	0.5998
mdb296	0.6526	0.0338	477.5393	37.9566	0.9831	12.323	136.5124	0.6662	0.6895	0.1157	0.1478	-0.7299	0.7413
mdb297	0.7694	0.0583	241.2228	23.9807	0.9709	9.8082	87.4395	0.4731	0.5134	0.1093	0.2221	-0.4107	0.4838
mdb298	0.4545	0.0933	324.2361	19.1088	0.9533	8.7204	62.0688	0.8948	0.9595	0.1008	0.3102	-0.5317	0.7078
mdb299	0.5787	0.1005	296.5426	22.9608	0.9497	9.5816	78.6343	0.7457	0.8154	0.0992	0.3263	-0.4094	0.5861
mdb300	0.5902	0.0595	308.3506	22.8504	0.9703	9.5515	78.5167	0.7275	0.7687	0.109	0.2255	-0.604	0.6979
mdb301	0.4672	0.0814	365.1108	21.6362	0.9593	9.29	71.6791	0.8731	0.9293	0.1036	0.2821	-0.571	0.7248
mdb302	0.8406	0.0424	225.9355	24.2913	0.9788	9.8763	90.8595	0.3546	0.384	0.1134	0.1755	-0.4014	0.4189
mdb303	0.9238	0.0168	216.1137	25.0353	0.9916	10.0323	96.5268	0.2128	0.2244	0.1203	0.0852	-0.5397	0.3918
mdb304	0.965	0.0139	204.0159	24.749	0.993	9.9787	97.6407	0.0988	0.1084	0.1211	0.0733	-0.1847	0.1481
mdb305	0.954	0.0124	208.4099	24.9073	0.9938	10.0095	97.456	0.1404	0.149	0.1215	0.0667	-0.4758	0.2995
mdb306	0.434	0.1084	372.2703	21.4955	0.9459	9.2501	70.0402	0.931	1.0065	0.0974	0.3432	-0.4885	0.6919
mdb307	0.9561	0.0093	209.708	25.037	0.9953	10.0352	98.3185	0.1226	0.1291	0.1224	0.0528	-0.5521	0.3062
mdb308	0.8089	0.0227	258.4556	25.8869	0.9887	10.1908	96.1966	0.4	0.4157	0.1187	0.1083	-0.6962	0.6004
mdb309	0.9805	0.0043	204.0976	24.9304	0.9979	10.0154	99.0958	0.062	0.065	0.1238	0.0277	-0.5491	0.2219
mdb310	0.9778	0.0063	204.2692	24.9295	0.9968	10.0159	98.9559	0.0695	0.0739	0.1232	0.0382	-0.422	0.1988
mdb311	0.6769	0.0439	449.7449	38.0026	0.978	12.3383	138.5198	0.5908	0.6213	0.113	0.1803	-0.6324	0.6615
mdb312	0.7282	0.037	413.4198	37.5603	0.9815	12.2692	138.3619	0.5251	0.5507	0.1148	0.1582	-0.6412	0.6366
mdb313	0.8567	0.0233	238.282	25.5492	0.9884	10.1289	96.5123	0.3159	0.332	0.1185	0.1105	-0.6098	0.5032
mdb314	0.5592	0.0753	383.6721	27.4799	0.9623	10.4813	94.8608	0.7764	0.8283	0.1051	0.2673	-0.5445	0.6817
mdb315	0.7948	0.0144	487.1959	47.3277	0.9928	13.7813	179.1854	0.41	0.4185	0.121	0.0752	-0.7925	0.6583
mdb316	0.2918	0.0647	870.1009	25.784	0.9676	10.0501	77.7527	1.3773	1.4214	0.1077	0.2398	-0.7691	0.9121
mdb317	0.6039	0.0517	610.6246	45.8669	0.9742	13.5531	166.1317	0.6888	0.7246	0.111	0.2034	-0.6342	0.7009

Image No.	Energy	Contrast	Correlation	Sum of variances	Homog- -eneity	Sum Average	Sum Variance	Sum Entropy	Entropy	Diff variance	Diff Entropy	IMOC1	IMOC2
mdb318	0.4398	0.087	776.2783	43.5806	0.9565	13.1959	151.887	0.9072	0.9674	0.1023	0.2955	-0.5639	0.7298
mdb319	0.7999	0.0196	480.6984	47.4452	0.9902	13.7989	179.9048	0.3988	0.4118	0.1195	0.0965	-0.7256	0.6181
mdb320	0.9958	0.0007	393.6489	48.8076	0.9996	14.0035	195.6446	0.0164	0.0169	0.1248	0.0059	-0.6817	0.1317
mdb321	0.8473	0.0166	455.1998	47.8379	0.9917	13.8521	183.1976	0.3266	0.3379	0.1203	0.0843	-0.7155	0.5641
mdb322	0.7782	0.0218	494.2636	47.3905	0.9891	13.788	178.2727	0.4505	0.4652	0.1189	0.1049	-0.7293	0.6467

*Diff entropy = Difference Entropy
*Diff variance = Difference variance
*IMOC1 = Information measure of Correlation 1
*IMOC2 = Information measure of Correlation 2

Table 3 Extracted Proposed Features.

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb001	22494.25	22038.63	18961.63	14359.75	3659.375	8	0	22095	21616	9180	65	0.285714	0.0413
mdb002	22500	22500	22118.38	20794.5	18838.5	0	0	22500	22405	2030	62	0.267081	0.0106
mdb003	22500	22499.5	22187.75	20751.25	16382.63	0	0	22500	22332	3524	58	0.242236	0.0264
mdb004	22500	22500	22124.75	19968.5	12168.63	0	0	22500	22479	4803	56	0.229814	0.0275
mdb005	0	0	0	0	0	22500	2285	0	0	0	43	0.149068	0.1062
mdb006	2534.5	112.375	0	0	0	20017	16085	160	19	0	66	0.291925	0.0245
mdb007	5354.25	495.5	0	0	0	17253	11265	641	112	0	76	0.354037	0.0502
mdb008	8225.125	996.25	1	0	0	14425	9406	1175	381	5	73	0.335404	0.0716
mdb009	2806.125	77.125	27.5	23.625	9.625	19778	14371	98	41	11	92	0.453416	0.0274
mdb010	12298.13	4521	1273.25	122.125	0	10133	4170	4835	3241	1384	117	0.608696	0.073
mdb011	19190.63	2725.5	142.125	30.375	25.375	3511	870	3354	975	128	87	0.42236	0.0987
mdb012	2714.875	19	0	0	0	19882	14456	30	3	0	66	0.291925	0.0349
mdb013	22446.88	22249.88	18011.88	9373	1972.375	55	39	22276	21983	13599	88	0.428571	0.057
mdb014	22500	22070.88	19182.88	13512.5	6010.75	0	0	22170	21671	9749	63	0.273292	0.0427
mdb015	16214.13	4228	64.125	0	0	6442	3342	4765	2196	93	66	0.291925	0.0856
mdb016	731	83	0	0	0	21783	7834	95	20	0	76	0.354037	0.0686
mdb017	20523.63	14366.5	4896.375	570.625	0	2034	918	14891	11799	5345	69	0.310559	0.0439
mdb018	22477.38	19387.13	6089.75	300.875	0	29	0	19749	16707	7061	49	0.186335	0.0171
mdb019	11971.88	1291.75	9.125	0	0	10751	4638	1518	589	20	93	0.459627	0.094
mdb020	3497.75	323.125	0	0	0	19000	13107	395	154	0	91	0.447205	0.0323
mdb021	22005.13	17315.75	5287.625	734.375	13.375	522	138	17861	14258	5858	84	0.403727	0.0339
mdb022	21457.38	18370.63	12663.63	6185.625	820.125	1094	534	18534	17387	10116	81	0.385093	0.0705
mdb023	22403.75	21553.13	12491.25	6074.375	987.875	98	38	21716	19912	10187	76	0.354037	0.0448
mdb024	18263.38	9364.25	3252.875	1021.875	17.875	4329	2028	9813	7428	3262	94	0.465839	0.0652
mdb025	11.625	0	0	0	0	15536	1995	0	0	0	84	0.403727	0.0696
mdb026	175.25	3	0	0	0	17591	6161	5	0	0	94	0.465839	0.0484
mdb027	0	0	0	0	0	11795	2075	0	0	0	80	0.378882	0.0611

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb028	10.625	0	0	0	0	10334	416	0	0	0	89	0.434783	0.0408
mdb029	22500	22500	22166.13	17726.13	6928.25	0	0	22500	22471	9036	54	0.217391	0.047
mdb030	22500	22490.13	21886.5	17630.5	5359.875	0	0	22496	22446	10039	53	0.21118	0.0677
mdb031	21171.5	18710.38	12532	6466.125	1284	1357	661	19004	17236	9665	102	0.515528	0.0605
mdb032	20856.38	16313.25	9027.5	3549.375	432.25	1696	778	16629	14134	7920	85	0.409938	0.0621
mdb033	19719.5	17346.63	11510	3188.375	4.25	2797	1256	17591	16049	11645	94	0.465839	0.0655
mdb034	10990.75	3459.375	108.5	32.75	3	11581	6775	3892	1641	159	119	0.621118	0.0421
mdb035	20255.38	15695.75	5232.375	402.5	0	2292	1081	16048	13677	5993	105	0.534161	0.0347
mdb036	19695.38	13780.13	6795.875	882.5	0	2830	1263	14207	11897	7282	104	0.52795	0.0422
mdb037	22500	22339.88	19369	14160.5	2991.875	0	0	22388	22009	9729	52	0.204969	0.0359
mdb038	22500	21090.88	2604.25	0	0	0	0	21491	18093	3672	37	0.111801	0.0001
mdb039	22447.25	21428.75	14690.88	6579.5	1296.75	60	5	21558	20335	11930	73	0.335404	0.0635
mdb040	20559.5	18384.38	14003.13	7468.5	767	1958	1169	18539	17620	11085	120	0.627329	0.0782
mdb041	13465.38	6354.875	1933.75	528.375	5.375	8883	3859	6779	4483	1968	121	0.63354	0.0579
mdb042	10716.63	6752.875	2195.75	287.25	0	11715	6128	7027	5230	2357	110	0.565217	0.0418
mdb043	20012.5	11559.5	3369.5	609.75	12.625	2563	1137	12234	8541	3551	89	0.434783	0.0503
mdb044	22329.13	16809.63	3789.25	1116.375	167.875	191	17	17426	12738	3696	64	0.279503	0.0293
mdb045	5364.125	845.125	1	0	0	16954	8192	1000	323	5	94	0.465839	0.0618
mdb046	107.625	0	0	0	0	18148	2904	0	0	0	93	0.459627	0.0593
mdb047	5363.125	1455.125	168.375	0	0	17017	8434	1606	870	202	115	0.596273	0.0557
mdb048	9080.125	2217.125	11.625	0	0	13535	6817	2510	1058	21	87	0.42236	0.0644
mdb049	22500	22097.25	17465.25	9615.5	4849.25	0	0	22228	20788	10953	62	0.267081	0.0409
mdb050	22466.75	15866.63	3471.25	661.5	8.5	37	2	16655	12055	3812	56	0.229814	0.0192
mdb051	20205.38	10334.13	1048.875	87.75	0	2384	689	11033	6950	1300	74	0.341615	0.0472
mdb052	13670.63	5107.125	344.125	0	0	8949	4892	5494	3129	449	80	0.378882	0.06
mdb053	22306.5	21542.25	18328.75	10080	1147.375	199	116	21612	20894	13660	90	0.440994	0.0749
mdb054	22500	22429.75	17118	2915.25	0	0	0	22453	22021	17791	37	0.111801	0.0565
mdb055	22351.25	19573.13	11714.25	5392.125	1558.25	158	28	19901	17671	9173	72	0.329193	0.0478
mdb056	9644.75	3384.75	287.25	33.25	0	12975	7237	3684	2238	357	99	0.496894	0.0598

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb057	15107.88	5092	485	82.375	47	7542	4419	5602	3038	530	101	0.509317	0.065
mdb058	21022.63	12377.5	2543.5	521.75	33	1553	370	13000	8971	2700	69	0.310559	0.0462
mdb059	0	0	0	0	0	17640	794	0	0	0	71	0.322981	0.0848
mdb060	0	0	0	0	0	10138	25	0	0	0	64	0.279503	0.051
mdb061	20868.75	13202	1485.75	5.25	0	1684	688	13983	9001	1770	88	0.428571	0.0277
mdb062	5788.5	2	0	0	0	16877	10791	2	1	0	69	0.310559	0.0597
mdb063	20512.88	17014.25	8961.75	3560.5	426.5	1927	624	17422	15005	8114	121	0.63354	0.0451
mdb064	21984.38	18090.63	10576.88	3879.75	374.25	534	214	18449	16212	9720	113	0.583851	0.0562
mdb065	18437.25	15515.75	12008.13	8375	3501.75	4030	1494	15683	14611	6035	126	0.664596	0.0496
mdb066	4837.5	126	0	0	0	16873	7550	191	28	0	104	0.52795	0.0631
mdb067	22368.38	21669.75	21126	19916.75	12496.63	135	54	21697	21552	3759	82	0.391304	0.024
mdb068	22296.75	21434.38	19280.5	15741	5586	210	77	21525	21033	8457	81	0.385093	0.0457
mdb069	1330.25	9.375	0	0	0	15504	3430	15	0	0	90	0.440994	0.0521
mdb070	9931.75	1337.125	15.875	0	0	12736	7406	1596	451	23	58	0.242236	0.0635
mdb071	20908.13	15984.13	8507.5	4683.125	2096.125	1643	623	16499	13378	5555	87	0.42236	0.0435
mdb072	20573.63	19043.25	14888.63	10360.63	6310	1506	674	19183	18148	6846	180	1	0.0409
mdb073	16158.75	4941.5	1234.5	68.75	0	6508	2862	5366	3412	1369	84	0.403727	0.0864
mdb074	12767.25	5354.375	719.625	205.25	40.5	9289	3289	5731	3827	777	131	0.695652	0.0689
mdb075	1853.125	217.125	36.5	22.5	17.375	20686	9221	250	92	18	104	0.52795	0.0292
mdb076	6468.125	964.875	19.125	0	0	16123	11834	1141	422	33	62	0.267081	0.046
mdb077	108.25	1	0	0	0	22395	7745	1	0	0	65	0.285714	0.0324
mdb078	4189	183.125	0	0	0	18392	12684	269	28	0	65	0.285714	0.0416
mdb079	15860.63	6998.25	1646.5	434	42.375	6717	4593	7528	4660	1666	86	0.416149	0.0409
mdb080	10071.38	1996.5	68.625	0	0	12596	7198	2232	1072	86	81	0.385093	0.0634
mdb081	18380.63	13497	7659.25	4991.75	3071.875	4204	2229	13858	11443	3798	108	0.552795	0.0466
mdb082	9617.125	2272.875	36.25	0	0	12985	6095	2607	960	65	84	0.403727	0.0699
mdb083	21923.38	14921.88	2631.625	404.375	4	605	162	15557	11128	2899	75	0.347826	0.0266
mdb084	17222.88	6060.875	2209.875	889.625	115.25	5436	1921	6525	4462	1940	74	0.341615	0.0778
mdb085	10185.5	4951.5	821.875	110.125	16.375	12394	8347	5207	3665	946	112	0.57764	0.045

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb086	8061	2524.625	91.25	0	0	14410	7940	2814	1265	121	97	0.484472	0.0598
mdb087	10194.88	1996.5	40.125	0	0	12460	7159	2312	837	54	80	0.378882	0.0753
mdb088	5219.75	682.5	14.75	0	0	17364	9567	800	289	20	88	0.428571	0.06
mdb089	22177.25	18959.63	11984.5	5460	1036.875	334	170	19311	16960	9979	82	0.391304	0.0645
mdb090	22500	22486.75	21830.63	18578.75	11083.38	0	0	22490	22444	6820	59	0.248447	0.0477
mdb091	3157.125	329.125	8.375	0	0	19092	7888	395	154	11	93	0.459627	0.0573
mdb092	1153.5	69.875	0	0	0	20137	5795	106	3	0	90	0.440994	0.0752
mdb093	4108	1282.75	129.375	0	0	17469	8122	1354	831	166	121	0.63354	0.0595
mdb094	77	2.25	0	0	0	21069	5668	4	0	0	90	0.440994	0.0567
mdb095	9198.375	2786.75	439.25	127.25	20.375	13406	9325	3016	1851	456	88	0.428571	0.045
mdb096	11363.88	5715.75	2343.125	930.875	155.125	11245	7810	5973	4555	1984	86	0.416149	0.0496
mdb097	21378.5	4869.25	4.125	0	0	1217	66	5908	1561	9	42	0.142857	0.06
mdb098	15234.38	899.375	0	0	0	7652	1226	1115	194	0	46	0.167702	0.1292
mdb099	20217.63	12179.75	2907.5	548.125	15.5	2355	1033	12847	8874	3112	78	0.36646	0.0486
mdb100	17123.5	6739.375	1412.25	303	2	5515	2953	7291	4534	1525	94	0.465839	0.08
mdb101	16210.75	9196.875	3279.75	629.75	9.125	6361	3294	9631	7087	3395	109	0.559006	0.0546
mdb102	13548	10221.13	4253.125	1476.625	124.375	7862	3775	10459	8708	4012	153	0.832298	0.0481
mdb103	9377.625	3536.75	527.375	2	0	11162	4846	3828	2339	641	115	0.596273	0.0542
mdb104	15584.63	10570.63	3079.75	502	5	6795	2885	10972	8227	3411	113	0.583851	0.0503
mdb105	22500	22500	22500	22500	22500	0	0	22500	2935	0	25	0.037267	0.0374
mdb106	22151	21128	17766.25	14478.38	9432.75	358	111	21201	20432	5468	117	0.608696	0.0303
mdb107	22500	22500	22145.75	19911.88	12087.75	0	0	22500	22472	5012	51	0.198758	0.0267
mdb108	22078	21207.75	16958.88	12076.88	8288.25	429	292	21327	20059	6987	100	0.503106	0.03
mdb109	21974.63	16820.63	7871	1534.75	8.375	558	77	17228	14918	8283	61	0.26087	0.0362
mdb110	17607.63	10940.38	1968.125	79	1	4952	2834	11515	7775	2206	88	0.428571	0.0335
mdb111	19202.63	6967.625	746.75	105.375	2	3412	1240	7686	3895	836	92	0.453416	0.0522
mdb112	10318	2206.875	28.875	0	0	12294	8141	2563	1033	49	70	0.31677	0.0532
mdb113	16653.5	6204.875	663.375	0	0	5975	3114	6739	3936	830	80	0.378882	0.0587
mdb114	19006	10348.5	198.375	0	0	3561	1797	11127	6242	320	82	0.391304	0.0296

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb115	16316.5	10435	3372.5	406.75	0	6245	3585	10756	8917	3761	91	0.447205	0.0477
mdb116	17129.13	13781.13	9407.875	7106.125	4984.875	5324	2055	13958	12711	3557	119	0.621118	0.0453
mdb117	14962.88	5289.375	1095.375	87.125	0	7669	4181	5725	3659	1236	91	0.447205	0.0664
mdb118	13629	9057.625	4000.625	1070.125	1	8910	4792	9396	7626	3940	94	0.465839	0.0499
mdb119	22063.88	16948.63	8213.75	2906	183.875	475	82	17451	14434	7609	66	0.291925	0.055
mdb120	21130.75	17885.88	8741.25	3295.75	194.75	1384	713	18187	15973	7998	132	0.701863	0.0512
mdb121	22438.13	19745.75	10897.63	5048.375	1310.625	68	10	19992	18112	8649	76	0.354037	0.0431
mdb122	18474.75	11036.75	6136	3198.75	718.125	4111	2171	11473	9439	4736	100	0.503106	0.0634
mdb123	21368	10916.25	2194.25	3	0	1212	227	11605	7945	2559	59	0.248447	0.0481
mdb124	22500	22211.5	18203.38	14065.88	8263.125	0	0	22288	21450	7244	60	0.254658	0.0349
mdb125	20429.5	16521.88	10871.5	6559.5	1668.875	2113	861	16839	14958	6996	110	0.565217	0.0569
mdb126	21256.13	15746	7348.875	2532.375	339.125	1291	530	16226	13341	6626	76	0.354037	0.0501
mdb127	14580.13	6450.625	669.75	16.75	0	8028	4842	6954	4315	797	72	0.329193	0.0471
mdb128	10311.75	4686.75	294.875	0	0	12232	5300	5090	2872	393	94	0.465839	0.0564
mdb129	22419.25	21449.5	11415.5	3153.125	84	85	16	21616	19923	11324	64	0.279503	0.0397
mdb130	22500	22335.13	18677	12389.5	4175	0	0	22369	21994	10675	61	0.26087	0.0488
mdb131	23.375	0	0	0	0	22477	7547	0	0	0	52	0.204969	0.0219
mdb132	284.75	0	0	0	0	22233	19838	0	0	0	37	0.111801	0.0018
mdb133	373.5	0	0	0	0	22142	18226	0	0	0	41	0.136646	0.0039
mdb134	21.625	0	0	0	0	22479	6026	0	0	0	50	0.192547	0.0861
mdb135	22464.88	4599	0	0	0	49	0	6134	868	0	32	0.080745	0.0216
mdb136	11057.5	85	0	0	0	11744	2797	125	13	0	42	0.142857	0.103
mdb137	21076.38	19653.88	13452.5	6375	784.75	1440	921	19740	18825	10680	87	0.42236	0.0496
mdb138	19575	17101.13	14350.88	11716.38	9207.125	2950	1314	17216	16412	4090	108	0.552795	0.0366
mdb139	12376.88	709.875	6.25	0	0	10397	2581	940	155	10	70	0.31677	0.0918
mdb140	1853	46.25	0	0	0	20717	16856	58	17	0	48	0.180124	0.0199
mdb141	1047.75	1	0	0	0	21487	18298	9	0	0	50	0.192547	0.0103
mdb142	3671	128.75	0	0	0	18930	13615	161	48	0	46	0.167702	0.0322
mdb143	12858.88	1326	2	0	0	9856	3948	1661	413	8	47	0.173913	0.0915

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb144	7514.25	872.125	4	0	0	15121	10633	1047	428	8	61	0.26087	0.0533
mdb145	17777.38	12744.63	8597.125	5451.625	1977.125	4797	2875	13052	11103	5242	91	0.447205	0.0659
mdb146	20952.63	17993.88	10876.75	5927.625	2234.625	1569	1059	18261	16428	7417	90	0.440994	0.0485
mdb147	4411.875	534.5	137.125	79.5	29	18188	11847	597	286	84	96	0.478261	0.0461
mdb148	11710.38	4952.625	1260.75	134.25	56.25	10917	6453	5255	3638	1390	106	0.540373	0.0582
mdb149	2824.875	409.5	0	0	0	19713	7126	479	141	0	78	0.36646	0.061
mdb150	354.875	0	0	0	0	21656	5645	0	0	0	71	0.322981	0.053
mdb151	22500	22499	15855.38	3618.625	0	0	0	22499	22475	16527	41	0.136646	0.0643
mdb152	22500	22500	22476	13732.25	149.25	0	0	22500	22500	19027	28	0.055901	0.1146
mdb153	5	2	0	0	0	4573	5	2	2	0	101	0.509317	0.021
mdb154	13.5	0	0	0	0	22431	10463	0	0	0	58	0.242236	0.0378
mdb155	20138.13	4369.125	71.625	0	0	2514	418	5214	1511	92	55	0.223602	0.0791
mdb156	13112.25	988.75	0	0	0	9589	4633	1347	129	0	68	0.304348	0.0876
mdb157	8698.375	347.375	0	0	0	14061	7569	439	113	0	58	0.242236	0.0823
mdb158	3452	5	0	0	0	19157	13317	10	0	0	48	0.180124	0.039
mdb159	21752.25	11383.63	1885.375	257.375	7.25	808	98	12333	7571	2033	74	0.341615	0.0375
mdb160	12303.75	3289.125	318.5	48.625	0	10374	6423	3616	1852	341	78	0.36646	0.072
mdb161	21330	16183.88	8345.875	3393	206	1206	506	16673	13740	7437	94	0.465839	0.0455
mdb162	21815.5	20283.13	16963.63	10747	3314.375	698	378	20407	19412	10677	102	0.515528	0.0631
mdb163	22488.88	21795.88	17114.5	9767.75	3599.75	13	0	21891	21060	11359	75	0.347826	0.0514
mdb164	22500	22500	22363.75	18703.13	7235.25	0	0	22500	22500	9406	41	0.136646	0.0583
mdb165	20334.63	16298	6666.625	245.25	2	2196	1146	16642	14424	7466	102	0.515528	0.0236
mdb166	20766.13	16945.25	7639	1187.375	3.125	1756	1227	17282	14995	8113	83	0.397516	0.0378
mdb167	62.75	0	0	0	0	22433	1706	0	0	0	61	0.26087	0.1047
mdb168	34.25	0	0	0	0	22458	4950	0	0	0	64	0.279503	0.0827
mdb169	19317.13	11564.13	3697.625	809.875	2	3288	1521	12027	9396	3917	76	0.354037	0.0583
mdb170	11459.13	5945.5	1585.5	276.25	36.125	11115	6531	6303	4349	1735	114	0.590062	0.0502
mdb171	22500	22500	22500	22500	22500	0	0	22500	7584	0	19	0	0.0028
mdb172	22500	22500	22497.5	22242.25	21850	0	0	22500	18809	459	47	0.173913	0.0029

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb173	11545.5	1223.125	9.625	0	0	11131	4851	1552	392	13	55	0.223602	0.0636
mdb174	7177.875	585.625	0	0	0	15429	11009	745	143	0	47	0.173913	0.043
mdb175	21443.75	13945.88	2699.875	112.25	0	1118	308	14533	10643	3190	71	0.322981	0.0376
mdb176	21485	17833.75	8878.875	359.625	0	1036	632	18225	15914	9665	73	0.335404	0.0267
mdb177	22003.13	18602.13	11395	4981.625	956.125	522	105	18917	16849	9150	76	0.354037	0.0456
mdb178	21797.75	16459.75	7611.5	2772.75	515.5	741	236	16903	14024	6644	73	0.335404	0.0483
mdb179	22500	22500	22500	22500	22433.5	0	0	22500	4659	0	34	0.093168	0.0698
mdb180	22500	22500	22500	22283.63	21121.63	0	0	22500	13226	545	47	0.173913	0.0104
mdb181	10090.38	3005.875	199.875	0	0	12513	7655	3341	1682	251	80	0.378882	0.0455
mdb182	8352.25	2103.625	113.375	1	0	14252	9075	2313	1248	133	84	0.403727	0.0493
mdb183	5185.375	384.875	14.75	0	0	17423	12567	453	214	20	65	0.285714	0.0418
mdb184	9778	447.75	0	0	0	12930	6843	577	86	0	50	0.192547	0.0809
mdb185	10220.5	2033.5	2	0	0	12404	6065	2344	761	2	95	0.47205	0.0911
mdb186	14254.75	6180	256.75	110.875	59.25	8332	4461	6720	3654	246	119	0.621118	0.0563
mdb187	20010.25	14685.13	6166.375	615.625	8.5	2554	1288	15076	12607	6701	91	0.447205	0.0393
mdb188	18439.63	14618.13	8352.125	2423.25	47	4096	2261	14956	12995	8078	91	0.447205	0.0505
mdb189	1	0	0	0	0	14248	1281	0	0	0	99	0.496894	0.0544
mdb190	0	0	0	0	0	188	0	0	0	0	60	0.254658	0.1187
mdb191	22500	22500	22442.88	18452.25	8645.125	0	0	22500	22500	8366	49	0.186335	0.0591
mdb192	21964	16398.63	4931.25	637.375	0	559	287	17139	12615	5310	72	0.329193	0.0288
mdb193	22500	22436.75	20642.13	18051.63	10888.75	0	0	22453	22125	4977	58	0.242236	0.0212
mdb194	22500	22500	20423.75	14495.88	5131.5	0	0	22500	22486	10744	47	0.173913	0.0468
mdb195	15175.75	7320.5	1601.5	90.5	0	7437	4389	7763	5176	1808	73	0.335404	0.0522
mdb196	7554	3347.75	1232.875	432.75	23.375	15045	8563	3473	2821	1163	89	0.434783	0.0404
mdb197	22500	22378.5	20193.88	16844.25	9607.625	0	0	22406	22021	6575	70	0.31677	0.0435
mdb198	22242.25	16733.88	3559.875	202.75	0	276	47	17396	12910	4093	59	0.248447	0.0187
mdb199	21933.88	17238.38	6131.875	1435.5	125.375	599	252	17752	14027	6209	84	0.403727	0.0379
mdb200	16695	12871.63	6047	2022.625	77.875	5836	3468	13138	11196	5927	117	0.608696	0.054
mdb201	20986.5	11656	2928.75	716.5	7.375	1570	626	12322	8249	3014	72	0.329193	0.0465

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb202	18882	16424.75	9375.75	2918.625	219.5	3644	2103	16634	14987	8947	106	0.540373	0.0563
mdb203	14075.13	2381.125	68.375	0	0	8603	4231	2858	1104	94	65	0.285714	0.0725
mdb204	10269.13	582.125	0	0	0	12213	5751	785	139	0	99	0.496894	0.0744
mdb205	81.125	0	0	0	0	22423	10152	0	0	0	55	0.223602	0.048
mdb206	288.375	1	1	0	0	22174	6421	1	1	1	89	0.434783	0.0792
mdb207	22320	21798.13	18122.63	8982.25	2367.875	187	40	21853	21295	13494	73	0.335404	0.0634
mdb208	22411.25	21768.38	17992.75	11705.25	2549.5	93	33	21844	21192	11266	78	0.36646	0.0524
mdb209	79.5	0	0	0	0	16732	4199	0	0	0	126	0.664596	0.0733
mdb210	27.875	0	0	0	0	19244	4645	0	0	0	110	0.565217	0.067
mdb211	22488.25	21744.5	18101.88	12565.13	4831.375	13	0	21822	21089	9585	70	0.31677	0.06
mdb212	21560.63	20541.13	16529.5	10814.25	4695.125	951	535	20630	20013	9043	100	0.503106	0.0447
mdb213	9720.375	1005.375	1	0	0	12926	5425	1259	362	3	83	0.397516	0.0821
mdb214	12694.63	3185.25	118.125	0	0	9865	4076	3558	1791	179	123	0.645963	0.0712
mdb215	22466.25	21259.63	15391.88	8521.25	2714.875	38	1	21409	20242	10588	68	0.304348	0.0474
mdb216	22500	22358.38	21725.13	19692.13	12902.13	0	0	22368	22151	4742	74	0.341615	0.0265
mdb217	22279.63	19816.63	14301	8485.5	2123.375	232	118	20009	18776	9601	103	0.521739	0.0564
mdb218	21067	17579.38	14085	11657.5	6605.75	1481	581	17815	16586	4850	91	0.447205	0.0484
mdb219	19548.13	10240.88	878.625	134.625	0	3022	1651	11087	5993	997	71	0.322981	0.0381
mdb220	12915	5516.5	1663.5	399	33	9699	5782	5866	4141	1709	95	0.47205	0.0655
mdb221	21724.5	20259.75	15266.63	9864.5	1864	797	386	20388	19556	10238	106	0.540373	0.0631
mdb222	21857.13	19895.75	15250.63	9419.875	2133.75	664	326	20131	18651	9546	82	0.391304	0.0545
mdb223	16309.88	7914.25	733.5	1	0	6274	3882	8623	4638	936	79	0.372671	0.0409
mdb224	14705.75	7351.375	1550.875	109.125	0	7347	3338	7748	5513	1801	143	0.770186	0.0536
mdb225	17782.25	13072.75	5231.75	968.125	23.25	4771	2318	13428	11176	5466	105	0.534161	0.0451
mdb226	21697.63	16155.88	5388.125	711.625	4	837	323	16690	13255	5865	70	0.31677	0.0334
mdb227	21731.25	14652.13	3049.375	323	67.25	809	320	15454	10723	3364	86	0.416149	0.0289
mdb228	18021.5	11173.5	1597.5	50.25	0	4546	2301	11753	7756	1888	94	0.465839	0.043
mdb229	12916.75	5417	1727.5	708.125	16.25	9705	5652	5744	4038	1451	85	0.409938	0.0609
mdb230	15337.13	7737	1571.875	148.25	0	7254	4240	8200	5677	1796	88	0.428571	0.0418

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb231	0	0	0	0	0	19394	1	0	0	0	41	0.136646	0.0229
mdb232	5	0	0	0	0	22350	2536	0	0	0	60	0.254658	0.0791
mdb233	14619.75	9349.625	1661.375	360.375	11.25	7957	5302	9749	6951	1860	92	0.453416	0.0449
mdb234	10664.13	5754	606.625	30	0	8047	1969	6165	3828	744	149	0.807453	0.0523
mdb235	13631.25	6338.625	748.375	207.5	0	8939	4775	6817	4001	815	100	0.503106	0.0522
mdb236	22499.25	22250.25	19118	11515.63	4255.75	1	0	22274	22009	11893	65	0.285714	0.0562
mdb237	7961.375	1000.75	0	0	0	14665	10471	1215	283	1	66	0.291925	0.0472
mdb238	9229.5	1085.25	12.75	0	0	13441	7770	1223	580	21	71	0.322981	0.0678
mdb239	22500	22492.63	21904.25	20223.63	13420	0	0	22494	22465	4412	54	0.217391	0.0234
mdb240	22500	22500	22500	22460.38	21854.38	0	0	22500	22494	184	35	0.099379	0.0013
mdb241	22500	22380.25	18615.88	8824	459.625	0	0	22406	22101	16205	57	0.236025	0.0701
mdb242	20690.13	15725.63	7765.75	2354.375	4	1844	921	16227	13132	7808	75	0.347826	0.0528
mdb243	22500	22499.5	22439.25	21841.38	16177.38	0	0	22499	22477	2425	62	0.267081	0.0111
mdb244	22500	22500	21635.5	18436.88	10057.25	0	0	22500	22491	7222	57	0.236025	0.0447
mdb245	8039.5	1038.75	22.375	0	0	14612	9481	1211	450	26	101	0.509317	0.0562
mdb246	12118.25	4783.625	259.5	0	0	7417	2339	5149	3048	342	142	0.763975	0.0694
mdb247	7636.125	612.875	0	0	0	15044	9568	735	188	0	63	0.273292	0.0664
mdb248	5657.25	343.875	3.125	0	0	16994	11973	406	139	4	64	0.279503	0.0593
mdb249	20534.38	11798.25	2622.125	614.375	25.25	2043	822	12438	8650	2817	75	0.347826	0.0504
mdb250	19579.88	11332.75	2958.375	814.125	4	3003	1429	11905	8418	3040	76	0.354037	0.0576
mdb251	103	0	0	0	0	21763	6503	0	0	0	73	0.335404	0.0682
mdb252	7264.125	670.5	22.75	4.25	1	15399	9921	818	241	20	83	0.397516	0.0608
mdb253	22500	22500	22497.5	22075.75	15326.13	0	0	22500	22500	2644	35	0.099379	0.0143
mdb254	22500	22500	21543.88	13830.63	2640.375	0	0	22500	22489	13948	43	0.149068	0.0753
mdb255	12975.13	617.875	0	0	0	9892	1565	767	157	0	41	0.136646	0.1045
mdb256	18021.88	355.125	1	0	0	4854	659	478	57	1	50	0.192547	0.1754
mdb257	22500	22499.25	21925	19270	14138.13	0	0	22499	21294	4750	67	0.298137	0.0202
mdb258	22500	22500	22499.5	22178.63	20599	0	0	22500	17352	997	62	0.267081	0.0201
mdb259	22500	22500	22500	22469.75	20828.63	0	0	22500	11299	220	45	0.161491	0.0443

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb260	22500	22500	22500	22487.25	22467.25	0	0	22500	5459	20	44	0.15528	0.0545
mdb261	22070.5	19834.88	16954.13	14155.5	7931.75	440	241	20005	19202	6133	83	0.397516	0.0386
mdb262	22500	22500	21410	17149.88	7236.625	0	0	22500	22499	7985	42	0.142857	0.0343
mdb263	6718.375	1066.125	0	0	0	15856	7477	1239	432	0	81	0.385093	0.0671
mdb264	681.125	9.375	0	0	0	14337	2638	17	0	0	120	0.627329	0.0651
mdb265	22304.63	21181.25	13390.25	3777.875	550.375	204	37	21281	20424	12574	67	0.298137	0.045
mdb266	19180.25	9671.75	1432	79	0	3417	1606	10378	6115	1656	84	0.403727	0.0482
mdb267	21086.38	15135	6200.625	909.625	1	1469	271	15618	12583	6589	62	0.267081	0.0474
mdb268	22500	22254.38	13938.5	3737.875	380.75	0	0	22303	21770	13352	63	0.273292	0.0409
mdb269	21701.88	16815.13	5632.5	853.5	1	824	353	17294	13873	6152	66	0.291925	0.0391
mdb270	17084.5	6754.75	1353.5	155.375	11.25	5551	2484	7164	4714	1493	102	0.515528	0.0666
mdb271	0	0	0	0	0	22500	1149	0	0	0	42	0.142857	0.129
mdb272	0	0	0	0	0	22500	1841	0	0	0	46	0.167702	0.0921
mdb273	708.125	50.625	0	0	0	21809	8694	60	12	0	64	0.279503	0.0486
mdb274	956.125	47.375	0	0	0	20758	3168	49	17	0	88	0.428571	0.0514
mdb275	18182.25	10555.88	3106.625	599.5	5.875	4417	2420	10987	8390	3354	78	0.36646	0.0539
mdb276	15080.88	4904.25	467.75	0	0	7545	4236	5431	2854	604	72	0.329193	0.0645
mdb277	16223.25	5922.375	813.75	20.875	0	6398	2945	6481	3922	996	84	0.403727	0.0593
mdb278	13767.13	5170.625	624.5	28.125	0	8855	4190	5567	3362	745	90	0.440994	0.0642
mdb279	22500	21953.63	14867	5308	1402	0	0	22046	21246	12698	67	0.298137	0.0447
mdb280	22495	18841.25	8973	5136.375	2861.75	7	0	19417	15815	5538	73	0.335404	0.0264
mdb281	22440.38	22053.5	18388.88	10890.13	2189.625	63	18	22109	21586	12621	72	0.329193	0.0638
mdb282	21557.5	18342.75	10782.13	3395.5	20.5	966	473	18696	16366	10628	70	0.31677	0.0592
mdb283	21888.13	16822.88	7428.75	3179	200.375	642	174	17423	13864	6208	69	0.310559	0.0399
mdb284	20128	15484.75	4086.875	750.5	22.375	2422	1449	15915	12294	4353	97	0.484472	0.0312
mdb285	20232.25	14570.75	3738	319.125	0	2317	1195	14970	12149	4298	92	0.453416	0.0357
mdb286	8850.25	2142.25	429	13.375	0	13793	8330	2343	1403	469	87	0.42236	0.0607
mdb287	20716.13	18374.38	14891.13	9491.625	1714.625	1808	927	18507	17673	10008	111	0.571429	0.068
mdb288	22442	22152.75	21037.38	19219	14884.88	60	24	22186	19742	3657	99	0.496894	0.02

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb289	21720.75	18522.63	12617.75	6773.125	2068.375	816	239	18804	17090	9284	78	0.36646	0.0529
mdb290	21809.63	20493.5	14449.38	6159.5	90	702	396	20637	19709	13455	77	0.360248	0.06
mdb291	22500	22291.63	21007.5	18084.63	10234.25	0	0	22307	22104	6144	65	0.285714	0.0328
mdb292	21437.63	18455.13	14285.38	9826.125	4046	1085	542	18671	17376	7662	91	0.447205	0.0486
mdb293	4990.375	786.125	67	27.875	10.625	17601	8269	884	377	49	106	0.540373	0.0687
mdb294	1614.375	101.875	0	0	0	20932	15254	122	27	0	60	0.254658	0.0148
mdb295	22500	22498.5	22019.25	20109.88	10031.25	0	0	22500	22472	5904	52	0.204969	0.0318
mdb296	22418.13	18383.63	8574.625	4042.75	1250.125	97	0	18827	15842	6569	66	0.291925	0.0338
mdb297	64.5	0	0	0	0	22439	6764	0	0	0	54	0.217391	0.0583
mdb298	1	0	0	0	0	22417	1265	0	0	0	56	0.229814	0.0933
mdb299	1	0	0	0	0	22499	3534	0	0	0	48	0.180124	0.1005
mdb300	144.125	0	0	0	0	22353	7299	0	0	0	66	0.291925	0.0595
mdb301	12.625	0	0	0	0	22480	4126	0	0	0	58	0.242236	0.0814
mdb302	51.625	0	0	0	0	22452	7122	0	0	0	62	0.267081	0.0424
mdb303	945.75	43	0	0	0	21584	12245	55	11	0	62	0.267081	0.0168
mdb304	30.625	0	0	0	0	22471	9601	0	0	0	43	0.149068	0.0139
mdb305	437.625	61.5	0	0	0	22077	12447	72	29	0	63	0.273292	0.0124
mdb306	15.25	0	0	0	0	22478	2361	0	0	0	63	0.273292	0.1084
mdb307	816.125	0	0	0	0	21706	17630	9	0	0	52	0.204969	0.0093
mdb308	2967.5	689.875	134.875	42	0	19592	15291	781	405	145	78	0.36646	0.0227
mdb309	518.75	0	0	0	0	22000	18491	0	0	0	42	0.142857	0.0043
mdb310	567.75	0	0	0	0	21957	18949	0	0	0	49	0.186335	0.0063
mdb311	22500	22302.13	11211	3873.625	58.75	0	0	22403	20655	10920	43	0.149068	0.0439
mdb312	22494.25	19605.88	9423.375	3099	311.875	7	0	20068	17193	8799	59	0.248447	0.037
mdb313	2688	262	0	0	0	19894	15345	297	107	0	55	0.223602	0.0233
mdb314	8814.125	1172.5	58.25	0	0	13916	6608	1293	645	72	79	0.372671	0.0753
mdb315	22456.88	22285.25	21443.5	20061.75	15828.75	44	19	22304	22178	2639	84	0.403727	0.0144
mdb316	8422.375	2242.375	35	0	0	12795	3605	2583	1012	48	122	0.639752	0.0647
mdb317	22500	22500	22026	17609.5	1254.75	0	0	22500	22500	12706	44	0.15528	0.0517

Image No.	WA>160	WA>175	WA>190	WA>200	WA>210	PC 110-160	PC 140-155	PC>175	PC 180-230	PC 190-205	Contrast	Normalized Contrast	GLCM Contrast
mdb318	22500	22500	22403.75	13600.75	746.375	0	0	22500	22464	16980	49	0.186335	0.087
mdb319	22500	22447.5	21947.25	20201	14750.75	0	0	22453	22318	3562	68	0.304348	0.0196
mdb320	22500	22500	22500	22500	22469.75	0	0	22500	18132	0	37	0.111801	0.0007
mdb321	22499.25	22339.63	22026.63	20867	15634.25	1	0	22348	21632	2891	75	0.347826	0.0166
mdb322	22500	22475.5	21800	20029	17198.25	0	0	22481	19796	3085	78	0.36646	0.0218

*WA = White Area
*PC = Pixel Count

6.2 CLASSIFICATION

The classifiers were applied on GLCM features, proposed features and the combination of both, separately, to check and compare their accuracy of classification for the same database. The results are presented in the following tables. Table 4 to 6 shows the results of: only GLCM features, three proposed features and combination of both (GLCM+proposed) respectively. As already discussed in the chapter 5, the value of *k* has large impact on the classification accuracy. For avoiding conflict odd k-values are taken in to consideration. The k-value that is showing best accuracy is finally chosen as the best result given by the classifier. It can be observed that the best accuracy was given by KNN classifier (K=5). One can also note one important thing here, as the value of k increases the accuracy also increases till a saturation point comes when accuracy becomes stable and does not increase further anymore.

Table 4 Results of Classifiers applied on GLCM features.

S.No.	**Classifier**	**Accuracy**
1.	Naïve Bayes	57.8 %
2.	Random Forest	56.9%
3.	KNN (K=1)	55%
4.	KNN (K=3)	67%
5.	KNN (K=5)	71.6%
6.	KNN (K=7)	71.6%
7.	KNN (K=9)	67.9%

Table 5 Results of Classifiers applied on the proposed features.

S.No.	Proposed Features	Accuracy
1.	Naïve Bayes	67%
2.	Random Forest	66.1%
3.	KNN (K=1)	66.1%
4.	KNN (K=3)	71.6%
5.	KNN (K=5)	72.5%
6.	KNN (K=7)	71.6%
7.	KNN (K=9)	67%

Table 6 Results of Classifiers applied on the "GLCM+proposed" features.

S.No.	Classifier	Accuracy
1.	Naïve Bayes	69.7%
2.	Random Forest	70.6%
3.	KNN (K=1)	67%
4.	KNN (K=3)	69.7%
5.	KNN (K=5)	70.6%
6.	KNN (K=7)	67.9%
7.	KNN (K=9)	66.1%

Here, we can see that every time *k*-value is changed, the accuracy is changing. Finding *k*-value that gives best result is of prime importance. From these tables we can see that initially when *k*=1, the accuracy is less. As *k*-value is increasing, the accuracy reached a pea*k* and then started decreasing for *k*=5 and 7. The maximum accuracy is found for *k*=5, and hence this value of *k* is used in further analysis.

6.2.1 Dimensionality Reduction

It has been mentioned previously, a large number of features make the data analysis more complex, hectic and takes more time; hence there is need to have feature selection process for dimensionality reduction. The features are decremented in such a way that, the maximum accuracy is yielded upon classification. In this work, 7 optimal features are extracted from 16 initial features using (Correlation Feature Selection) CFS. Here are the 7 optimal features which are selected after dimensionality reduction:

1) Pixel Count
2) White area
3) Contrast
4) Correlation
5) Sum average
6) Sum variance
7) Information measure of correlation 1.

The following table shows the results of classifier after feature selection process. The best result was produced by KNN at k-1.

Table 7 Results of classifier applied on the seven optimal features.

S.No.	**Classifier**	**Accuracy**
1.	Naïve Bayes	65.1%
2.	Random Forest	66.1%
3.	KNN (K=1)	71.6%
4.	KNN (K=3)	67.9%
5.	KNN (K=5)	68.8%
6.	KNN (K=7)	67%
7.	KNN (K=9)	65.1%

6.2.2 Comparison of the Features

The following table presents the comparison of the features/ different set of features on the basis of their accuracy which have been produced by the different classifiers. The comparison has been made only with the best accuracy provided by the particular classifier on different features or set of features.

Table 8 Comparison of the features on the basis of accuracy.

S.No.	Features	Classifier	Accuracy
1.	GLCM Features	KNN (k=5)	71.6%
2.	Proposed Features	KNN (k=5)	72.5%
3.	GLCM + Proposed	KNN (k=5)	70.6%
4.	Optimal Features	KNN (k=1)	71.6%

Here, it can be observed that the maximum accuracy has been produced by the proposed features instead of widely known GLCM features. So, it can be said that the proposed features have contributed in the classification of the mammograms a lot. Even the combination of proposed and GLCM features has produced less accuracy. Table 9 represents the classification results such as TP rate, FP rate, Precision, Recall, AUC etc. for the proposed features which gave the best accuracy of 72.5% at k-5 with KNN classifier

Table 9 Table showing classification results such as TP rate, FP rate, AUC, precision values.

Class	TP Rate	FP Rate	Precision	Recall	F-Measure	AUC
Glandular	0.472	0.151	0.607	0.472	0.531	0.748
Dense	0.837	0.242	0.692	0.837	0.758	0.838
Fatty	0.867	0.038	0.897	0.867	0.881	0.978
Weighted Average	0.725	0.156	0.720	0.725	0.717	0.847

6.2.3 Confusion Matrix

Confusion matrix is another way of representing the classifier result. It gives numbers of correctly and falsely classifier instances out of the total instances. Higher the values of the diagonal and lower the values of the other elements, better is the result. Table 10 shows the confusion matrix resulted from the proposed features. The diagonal elements are shown in shaded cells. It is seen from the table that, in some instances non-diagonal elements show very low or zero values – which is always expected from a good classifier.

Table 10 Confusion Matrix.

Glandular	**Dense**	**Fatty**	**Classified as**
17	16	3	Glandular
7	36	0	Dense
4	0	26	Fatty

6.3 CLASSIFICATION INTO FATTY AND DENSE

As it can be seen from the results of classification that the results are not much satisfactory, so it was decided to classify the mammograms into fatty and dense only. So, out of 322 total mammograms, 218 mammograms were used for classification task. The extracted features, the dimensionality reduction scheme used in this task were same as previously used in classification of mammograms into fatty, glandular and dense. For the classification of fatty and dense the same classifiers were applied on GLCM features, proposed features, combination of both and optimal features separately, to check and compare their accuracy of classification. The results are presented in the following tables. Table 11 to 13 shows the results of GLCM features, proposed features and combination of both respectively.

Table 11 Results of the classifiers applied on GLCM features.

S.No.	Classifier	Accuracy
1.	Naïve Bayes	92.2%
2.	Random Forest	89.4%
3.	KNN (K=1)	89.9%
4.	KNN (K=3)	90.4%
5.	KNN (K=5)	91.7%
6.	KNN (K=7)	91.7%
7.	KNN (K=9)	90.8%

Table 12 Results of the classifiers applied on proposed features.

S.No.	Classifier	Accuracy
1.	Naïve Bayes	92.7%
2.	Random Forest	90.4%
3.	KNN (K=1)	90.8%
4.	KNN (K=3)	91.7%
5.	KNN (K=5)	93.1%
6.	KNN (K=7)	93.1%
7.	KNN (K=9)	93.1%

Table 13 Results of the classifiers applied on “GLCM+proposed” features.

S.No.	Classifier	Accuracy
1.	Naïve Bayes	92.2%
2.	Random Forest	91.3%
3.	KNN (K=1)	93.1%
4.	KNN (K=3)	92.7%
5.	KNN (K=5)	93.6%
6.	KNN (K=7)	92.7%
7.	KNN (K=9)	92.2%

6.3.1 Dimensionality Reduction

It has been mentioned previously in Chapter 5, Section 5.4.1, a large number of features make the data analysis more complex, hectic and takes more time; hence there is need to have feature selection process for dimensionality reduction. The features are reduced in such a way that, the maximum accuracy is yielded upon classification. The same method of Correlation Feature Selection is used for feature selection. Table 14 shows the extracted features used for the final analysis of classification of the mammograms into fatty and dense. Table 15 shows the results of the classifiers applied on the optimal features that have been extracted.

CFS assumes that useful feature subsets contain features that are predictive of the class but uncorrelated with one another. CFS computes a heuristic measure of the “merit” of a feature subset from pair-wise feature correlations and a formula adapted from test theory. Heuristic search is used to traverse the space of feature subsets in reasonable time; the subset with the highest merit found during the search is reported.

Table 14 Extracted features for final analysis for classification into fatty and dense.

Image No.	Correlation	Sum Average	Sum Variance	Info measure of correlation 1	PC >175	White area>190	Contrast
mdb003	461.1922	13.8396	181.8052	-0.618	22500	22187.75	58
mdb004	494.3947	13.7711	177.8419	-0.676	22500	22124.75	56
mdb005	268.7504	9.69	82.0373	-0.3092	0	0	43
mdb006	248.1643	10.1505	95.8643	-0.6434	160	0	66
mdb009	238.9989	10.1327	96.2588	-0.5691	98	27.5	92
mdb010	659.4145	10.8027	94.6835	-0.6874	4835	1273.25	117
mdb011	504.4882	11.4025	111.7591	-0.4933	3354	142.125	87
mdb012	238.1331	10.092	94.5451	-0.5192	30	0	66
mdb025	313.1126	8.5513	58.7371	-0.6401	0	0	84
mdb026	400.8128	9.0956	67.5986	-0.7352	5	0	94
mdb027	481.6605	8.0461	48.0846	-0.7626	0	0	80
mdb028	234.3413	7.8236	49.1292	-0.7557	0	0	89
mdb033	577.4558	11.9839	121.6989	-0.7014	17591	11510	94
mdb034	615.2796	10.7502	96.0354	-0.7859	3892	108.5	119
mdb035	398.4227	11.7514	124.3029	-0.7048	16048	5232.375	105
mdb036	442.0186	11.7487	121.3325	-0.7073	14207	6795.875	104
mdb037	758.9913	13.252	156.2105	-0.7669	22388	19369	52
mdb038	287.9001	11.9999	143.9758	0	21491	2604.25	37
mdb039	610.4149	12.564	138.421	-0.6284	21558	14690.875	73
mdb040	851.283	12.4514	128.8764	-0.6814	18539	14003.125	120
mdb053	787.4009	12.8587	142.521	-0.6321	21612	18328.75	90
mdb054	399.4392	12.2473	137.7597	-0.4848	22453	17118	37
mdb057	573.7744	11.1569	106.1006	-0.6537	5602	485	101
mdb058	390.1549	11.8085	125.769	-0.618	13000	2543.5	69
mdb059	274.6063	8.4981	58.7149	-0.5463	0	0	71
mdb060	177.7632	7.9327	53.3573	-0.601	0	0	64
mdb061	356.7433	11.7829	128.9401	-0.6621	13983	1485.75	88

Image No.	Correlation	Sum Average	Sum Variance	Info measure of correlation 1	PC >175	White area>190	Contrast
mdb062	284.0689	10.2574	94.6903	-0.4761	2	0	69
mdb063	572.1729	12.0503	123.9708	-0.7608	17422	8961.75	121
mdb064	495.2344	12.267	132.1637	-0.647	18449	10576.875	113
mdb065	1207.613	12.2493	121.7121	-0.8119	15683	12008.125	126
mdb066	413.6815	9.9868	81.5927	-0.6947	191	0	104
mdb067	496.8865	13.7481	176.5539	-0.7305	21697	21126	82
mdb068	711.5523	13.3564	157.8587	-0.7247	21525	19280.5	81
mdb069	369.6485	8.8326	62.9149	-0.7335	15	0	90
mdb070	449.9723	10.633	97.2577	-0.6287	1596	15.875	58
mdb075	236.0729	10.1065	95.4482	-0.5417	250	36.5	104
mdb076	365.9942	10.4523	96.1389	-0.664	1141	19.125	62
mdb077	245.1891	9.7977	87.7052	-0.6215	1	0	65
mdb078	295.1123	10.2456	93.731	-0.6331	269	0	65
mdb079	575.2735	11.3258	110.0368	-0.7602	7528	1646.5	86
mdb080	471.6404	10.6525	96.7245	-0.645	2232	68.625	81
mdb087	494.8573	10.6888	96.3868	-0.6078	2312	40.125	80
mdb088	360.2194	10.2583	89.6048	-0.6377	800	14.75	88
mdb091	419.7915	9.7502	77.5143	-0.7134	395	8.375	93
mdb092	426.0595	9.28	69.1399	-0.6517	106	0	90
mdb095	473.3187	10.6834	98.1997	-0.7235	3016	439.25	88
mdb096	600.9867	10.9461	100.3866	-0.7481	5973	2343.125	86
mdb097	376.8732	11.7281	125.1809	-0.489	5908	4.125	42
mdb098	535.1695	10.8462	98.2926	-0.4376	1115	0	46
mdb099	418.8618	11.743	122.7057	-0.6399	12847	2907.5	78
mdb100	542.8746	11.3787	110.1053	-0.5915	7291	1412.25	94
mdb101	598.3657	11.3286	107.6148	-0.7265	9631	3279.75	109
mdb102	883.6391	11.0147	95.9252	-0.8168	10459	4253.125	153
mdb103	794.8921	10.3148	84.3439	-0.7864	3828	527.375	115
mdb104	617.9309	11.226	104.5493	-0.7635	10972	3079.75	113
mdb105	1057.473	15.0485	203.2464	-0.7693	22500	22500	25

Image No.	Correlation	Sum Average	Sum Variance	Info measure of correlation 1	PC >175	White area>190	Contrast
mdb106	768.0389	13.2419	154.5883	-0.8089	21201	17766.25	117
mdb107	497.3111	13.7653	177.5824	-0.6934	22500	22145.75	51
mdb108	837.2252	13.0186	147.9757	-0.8216	21327	16958.875	100
mdb109	380.8508	12.0358	132.1784	-0.6513	17228	7871	61
mdb110	484.1734	11.4712	116.1801	-0.7566	11515	1968.125	88
mdb111	458.5503	11.5425	117.4476	-0.6485	7686	746.75	92
mdb112	498.9547	10.7432	99.112	-0.6872	2563	28.875	70
mdb125	843.5971	12.3093	125.8477	-0.7526	16839	10871.5	110
mdb126	476.0374	12.0395	126.9063	-0.6781	16226	7348.875	76
mdb129	420.265	12.2616	137.2704	-0.6406	21616	11415.5	64
mdb130	794.4651	13.0969	150.9313	-0.7172	22369	18677	61
mdb131	207.1514	9.9653	96.3333	-0.2237	0	0	52
mdb132	200.6251	10.0028	99.7125	-0.2318	0	0	37
mdb133	203.1046	10.0118	99.2284	-0.5033	0	0	41
mdb134	298.9194	9.5731	78.717	-0.4707	0	0	50
mdb135	301.182	11.9542	138.95	-0.3059	6134	0	32
mdb136	400.7479	10.5283	94.8681	-0.4501	125	0	42
mdb137	722.8799	12.4093	131.2444	-0.7509	19740	13452.5	87
mdb138	1039.372	12.7379	136.6169	-0.8347	17216	14350.875	108
mdb139	481.2895	10.7025	97.1286	-0.5346	940	6.25	70
mdb140	219.5046	10.0689	97.2674	-0.4905	58	0	48
mdb141	212.3851	10.0444	98.0803	-0.5834	9	0	50
mdb142	254.6537	10.1776	95.8863	-0.5885	161	0	46
mdb143	523.1649	10.8091	98.7542	-0.5486	1661	2	47
mdb144	376.1682	10.4746	96.0183	-0.6351	1047	4	61
mdb145	1012.214	11.9507	116.2356	-0.7452	13052	8597.125	91
mdb146	707.0019	12.3526	129.9468	-0.7529	18261	10876.75	90
mdb147	287.6807	10.2519	94.0991	-0.5923	597	137.125	96
mdb148	565.4255	10.8543	98.9666	-0.6969	5255	1260.75	106
mdb149	433.1623	9.6818	75.9233	-0.7063	479	0	78

Image No.	Correlation	Sum Average	Sum Variance	Info measure of correlation 1	PC >175	White area>190	Contrast
mdb150	367.4182	9.3191	72.715	-0.6872	0	0	71
mdb151	433.2077	12.3094	137.5262	-0.4984	22499	15855.375	41
mdb152	776.6091	13.1937	150.796	-0.4752	22500	22476	28
mdb153	135.5415	7.9855	60.6108	-0.4167	2	0	101
mdb154	309.2136	9.5345	79.2331	-0.7125	0	0	58
mdb155	441.229	11.5601	117.8554	-0.5054	5214	71.625	55
mdb156	553.5702	10.8859	99.681	-0.5736	1347	0	68
mdb157	362.3207	10.4447	94.7841	-0.4929	439	0	58
mdb158	246.1554	10.1528	95.735	-0.4885	10	0	48
mdb159	348.5254	11.8628	130.3948	-0.5715	12333	1885.375	74
mdb160	541.4336	10.8593	100.0012	-0.6267	3616	318.5	78
mdb161	514.0948	12.1479	128.7945	-0.7127	16673	8345.875	94
mdb162	866.3936	12.871	141.0624	-0.7048	20407	16963.625	102
mdb163	759.0802	12.8554	144.7358	-0.7057	21891	17114.5	75
mdb164	557.7499	13.6493	170.0516	-0.5566	22500	22363.75	41
mdb165	377.0275	11.7732	126.8679	-0.7535	16642	6666.625	102
mdb166	403.8534	11.9117	127.6144	-0.6904	17282	7639	83
mdb167	351.0268	8.8175	62.8116	-0.5094	0	0	61
mdb168	362.4898	9.3095	71.9699	-0.5646	0	0	64
mdb169	486.7077	11.6474	117.5197	-0.6532	12027	3697.625	76
mdb170	608.9825	10.8897	98.6917	-0.7492	6303	1585.5	114
mdb171	394.0976	14.0044	195.4112	-0.2345	22500	22500	19
mdb172	401.1099	13.9774	193.2744	-0.7398	22500	22497.5	47
mdb173	495.2284	10.7358	98.634	-0.6425	1552	9.625	55
mdb174	370.4234	10.462	96.3491	-0.6826	745	0	47
mdb179	837.7635	14.6185	191.2582	-0.6004	22500	22500	34
mdb180	408.7881	13.9972	191.9834	-0.5525	22500	22500	47
mdb183	301.0317	10.3014	95.4885	-0.6234	453	14.75	65
mdb184	416.7306	10.562	95.844	-0.5429	577	0	50
mdb193	582.3346	13.6047	170.2494	-0.8016	22453	20642.125	58

Image No.	Correlation	Sum Average	Sum Variance	Info measure of correlation 1	PC >175	White area>190	Contrast
mdb194	747.1211	13.2831	156.5228	-0.7131	22500	20423.75	47
mdb195	570.0302	11.1886	107.377	-0.702	7763	1601.5	73
mdb196	420.3048	10.5728	96.3001	-0.74	3473	1232.875	89
mdb197	641.5045	13.4969	164.3876	-0.6993	22406	20193.875	70
mdb198	310.0453	11.9655	137.7459	-0.5345	17396	3559.875	59
mdb199	376.6489	12.031	132.2224	-0.6263	17752	6131.875	84
mdb200	680.3331	11.5737	111.1613	-0.7508	13138	6047	117
mdb201	393.4652	11.8419	126.1704	-0.6202	12322	2928.75	72
mdb202	613.602	11.8912	119.2044	-0.7325	16634	9375.75	106
mdb203	566.0856	10.9807	102.4914	-0.6248	2858	68.375	65
mdb204	496.072	10.5221	91.7237	-0.649	785	0	99
mdb205	260.5852	9.726	84.7302	-0.565	0	0	55
mdb206	306.4047	9.5691	78.133	-0.5134	1	1	89
mdb207	749.1599	12.7629	141.0338	-0.6708	21853	18122.625	73
mdb208	809.1621	13.0196	147.7782	-0.7103	21844	17992.75	78
mdb215	712.0722	12.7447	142.5303	-0.7163	21409	15391.875	68
mdb216	505.7195	13.7488	176.6996	-0.7017	22368	21725.125	74
mdb221	859.7104	12.7752	138.6589	-0.708	20388	15266.625	106
mdb222	830.4803	12.75	139.0313	-0.7305	20131	15250.625	82
mdb223	527.0567	11.3276	112.0012	-0.7363	8623	733.5	79
mdb224	635.2901	11.0433	100.6532	-0.7523	7748	1550.875	143
mdb225	539.5393	11.5616	113.973	-0.7499	13428	5231.75	105
mdb226	352.4737	11.944	131.9402	-0.6114	16690	5388.125	70
mdb229	618.3682	11.0293	101.7053	-0.7065	5744	1727.5	85
mdb230	561.8914	11.2591	109.3665	-0.7464	8200	1571.875	88
mdb231	140.3723	8.0659	61.7817	-0.4213	0	0	41
mdb232	320.4558	8.7073	62.2959	-0.5798	0	0	60
mdb235	617.5876	11.0725	102.519	-0.7374	6817	748.375	100
mdb236	794.5095	13.0132	148.3163	-0.6877	22274	19118	65
mdb237	403.1913	10.5324	96.7942	-0.682	1215	0	66

Image No.	Correlation	Sum Average	Sum Variance	Info measure of correlation 1	PC >175	White area>190	Contrast
mdb238	429.1465	10.5815	96.2129	-0.6028	1223	12.75	71
mdb239	482.2966	13.7956	179.4565	-0.6919	22494	21904.25	54
mdb240	393.1296	13.9968	195.4291	-0.3855	22500	22500	35
mdb241	713.9442	12.7697	142.488	-0.6212	22406	18615.875	57
mdb242	481.2015	11.9907	125.4929	-0.6833	16227	7765.75	75
mdb243	416.3696	13.9401	189.5677	-0.6263	22499	22439.25	62
mdb244	568.1788	13.6302	169.8662	-0.6431	22500	21635.5	57
mdb245	391.1836	10.5069	96.1749	-0.6326	1211	22.375	101
mdb246	877.6816	10.3808	83.9034	-0.7708	5149	259.5	142
mdb247	362.1263	10.4439	95.2777	-0.5638	735	0	63
mdb248	298.6382	10.2957	94.8257	-0.5138	406	3.125	64
mdb249	411.4926	11.7762	123.6853	-0.6266	12438	2622.125	75
mdb250	467.3842	11.6954	119.1833	-0.6452	11905	2958.375	76
mdb251	352.6512	9.3652	73.5561	-0.611	0	0	73
mdb252	356.7914	10.4323	95.3863	-0.5858	818	22.75	83
mdb253	408.7961	13.9581	190.7998	-0.4513	22500	22497.5	35
mdb254	771.7734	13.2121	152.8702	-0.6036	22500	21543.875	43
mdb255	443.6568	10.62	95.7681	-0.4733	767	0	41
mdb256	567.8655	11.0047	100.2007	-0.3301	478	1	50
mdb257	531.0835	13.7134	174.7225	-0.7836	22499	21925	67
mdb258	432.1162	14.0235	189.4267	-0.5682	22500	22499.5	62
mdb259	714.1069	14.4658	190.2029	-0.6824	22500	22500	45
mdb260	684.8732	14.4339	189.4493	-0.6177	22500	22500	44
mdb261	791.0495	13.1969	152.3152	-0.7757	20005	16954.125	83
mdb262	628.9926	13.52	165.7913	-0.7352	22500	21410	42
mdb267	406.13	11.8485	125.5247	-0.636	15618	6200.625	62
mdb268	442.7973	12.3261	138.6192	-0.6437	22303	13938.5	63
mdb271	342.494	9.3972	73.0137	-0.3924	0	0	42
mdb272	371.1699	9.2525	70.589	-0.5393	0	0	46
mdb273	233.5433	9.9399	90.6251	-0.4375	60	0	64

Image No.	Correlation	Sum Average	Sum Variance	Info measure of correlation 1	PC >175	White area>190	Contrast
mdb274	369.6885	8.8561	63.8918	-0.7256	49	0	88
mdb281	793.4538	12.9469	145.78	-0.6641	22109	18388.875	72
mdb282	485.5008	12.1768	130.0016	-0.6391	18696	10782.125	70
mdb283	469.5721	12.1841	131.808	-0.7114	17423	7428.75	69
mdb284	422.1212	11.7826	124.2586	-0.7449	15915	4086.875	97
mdb285	392.3421	11.754	124.9542	-0.6796	14970	3738	92
mdb286	444.5058	10.5889	95.7839	-0.6461	2343	429	87
mdb287	953.8605	12.6408	133.3436	-0.7172	18507	14891.125	111
mdb288	528.8841	13.703	174.3004	-0.78	22186	21037.375	99
mdb289	712.7283	12.4802	133.2386	-0.7264	18804	12617.75	78
mdb290	638.8356	12.4633	134.1194	-0.681	20637	14449.375	77
mdb293	362.3064	10.1823	87.1666	-0.6171	884	67	106
mdb294	223.4128	10.0803	97.2393	-0.63	122	0	60
mdb295	490.7199	13.7784	177.9737	-0.6404	22500	22019.25	52
mdb296	477.5393	12.323	136.5124	-0.7299	18827	8574.625	66
mdb297	241.2228	9.8082	87.4395	-0.4107	0	0	54
mdb298	324.2361	8.7204	62.0688	-0.5317	0	0	56
mdb299	296.5426	9.5816	78.6343	-0.4094	0	0	48
mdb300	308.3506	9.5515	78.5167	-0.604	0	0	66
mdb301	365.1108	9.29	71.6791	-0.571	0	0	58
mdb302	225.9355	9.8763	90.8595	-0.4014	0	0	62
mdb303	216.1137	10.0323	96.5268	-0.5397	55	0	62
mdb304	204.0159	9.9787	97.6407	-0.1847	0	0	43
mdb305	208.4099	10.0095	97.456	-0.4758	72	0	63
mdb306	372.2703	9.2501	70.0402	-0.4885	0	0	63
mdb307	209.708	10.0352	98.3185	-0.5521	9	0	52
mdb308	258.4556	10.1908	96.1966	-0.6962	781	134.875	78
mdb309	204.0976	10.0154	99.0958	-0.5491	0	0	42
mdb310	204.2692	10.0159	98.9559	-0.422	0	0	49
mdb311	449.7449	12.3383	138.5198	-0.6324	22403	11211	43

Image No.	Correlation	Sum Average	Sum Variance	Info measure of correlation 1	PC >175	White area>190	Contrast
mdb312	413.4198	12.2692	138.3619	-0.6412	20068	9423.375	59
mdb313	238.282	10.1289	96.5123	-0.6098	297	0	55
mdb314	383.6721	10.4813	94.8608	-0.5445	1293	58.25	79
mdb315	487.1959	13.7813	179.1854	-0.7925	22304	21443.5	84
mdb316	870.1009	10.0501	77.7527	-0.7691	2583	35	122
mdb317	610.6246	13.5531	166.1317	-0.6342	22500	22026	44
mdb318	776.2783	13.1959	151.887	-0.5639	22500	22403.75	49
mdb319	480.6984	13.7989	179.9048	-0.7256	22453	21947.25	68
mdb320	393.6489	14.0035	195.6446	-0.6817	22500	22500	37
mdb321	455.1998	13.8521	183.1976	-0.7155	22348	22026.625	75
mdb322	494.2636	13.788	178.2727	-0.7293	22481	21800	78

*PC – Pixel Count

Table 15 Results of the classifier applied on the optimal features.

S.No.	Classifier	Accuracy
1.	Naïve Bayes	92.2%
2.	Random Forest	89.9%
3.	KNN (K=1)	90.8%
4.	KNN (K=3)	92.7%
5.	KNN (K=5)	94.5%
6.	KNN (K=7)	94%
7.	KNN (K=9)	91.7%

6.3.3 Confusion Matrix

For two classes the plot of what confusion matrix represents is shown in Table 16. Table 17 represents three parameters (accuracy, sensitivity and specificity) that can be derived from the contents of confusion matrix. Table 18 to 21 shows the confusion matrix of all the set of features – only GLCM features, three proposed features, GLCM+Proposed features and optimal features. The diagonal elements are shown in shaded cells.

Table 16 Representation of Confusion Matrix

Actual Class	Predicted Class	
	Dense	Fatty
Dense	True Positive(TP)	False Negative (FN)
Fatty	False Positive (FP)	True Negative (TN)

Table 17 Parameters derived from confusion matrix

Parameter	Formula
Accuracy	TP+TN/TP+TN+FP+FN
Sensitivity	TP/TP+FP
Specificity	TN/TN+FN

Table 18 Confusion matrix of Optimal Features.

Actual class	Predicted Class	
	Dense	Fatty
Dense	107	5
Fatty	7	99

Table 19 Confusion matrix of GLCM Features.

Actual class	Predicted Class	
	Dense	Fatty
Dense	106	6
Fatty	11	95

Table 20 Confusion matrix of Proposed Features.

Actual class	Predicted Class	
	Dense	Fatty
Dense	105	7
Fatty	8	98

Table 21 Confusion matrix of (GLCM+proposed) Features.

Actual class	Predicted Class	
	Dense	Fatty
Dense	106	6
Fatty	8	98

6.3.2 Comparison of the Accuracy, Sensitivity and Specificity

The following figures present the comparison of the features/ different set of features on the basis of their accuracy, sensitivity and specificity which have been produced by the different classifiers.

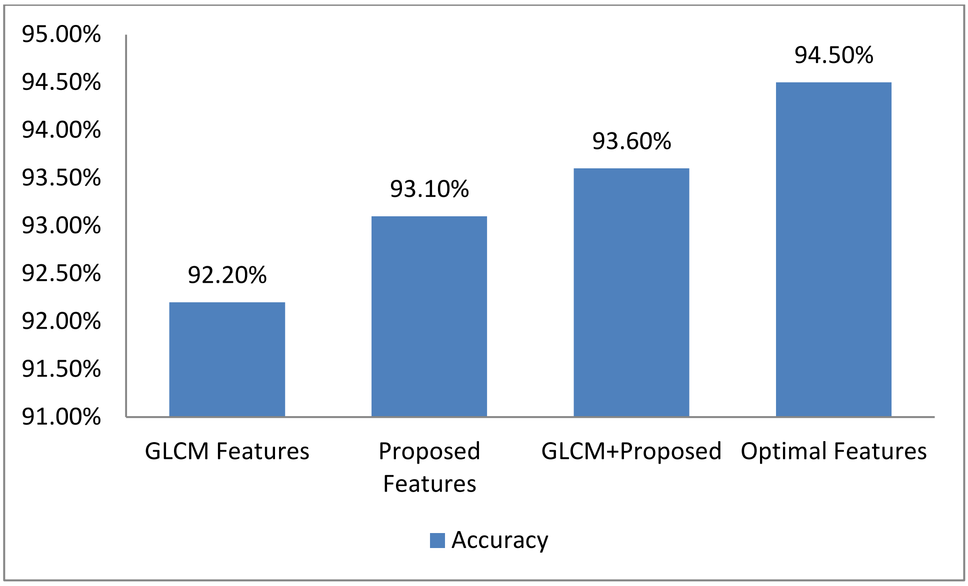

Figure 27 Comparison of accuracy for different set of features.

Here, it can be observed that the maximum accuracy has been produced by the optimal features instead of widely known only GLCM features, three proposed features and the combination of (GLCM+proposed) features.

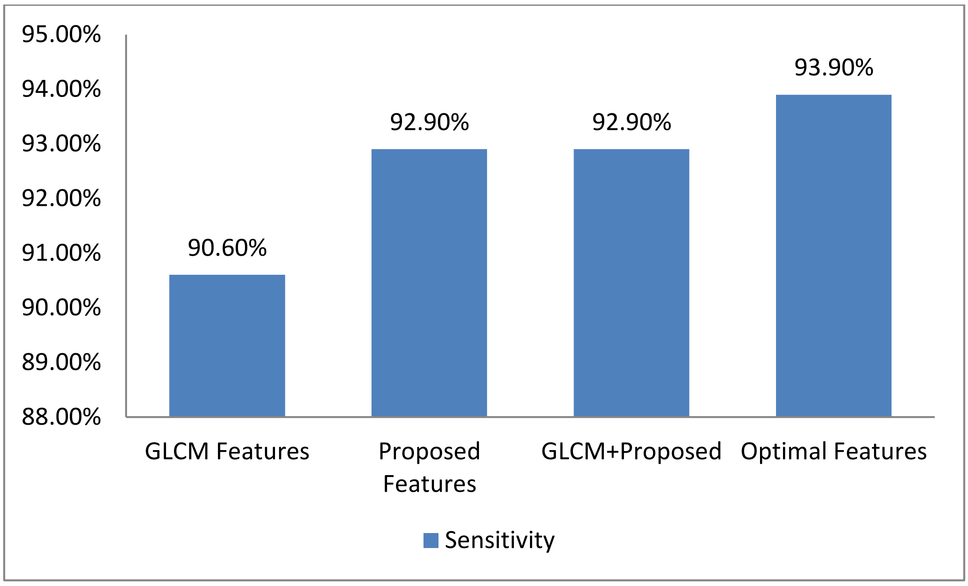

Figure 28 Comparison of sensitivity for different set of features.

Here, it can be observed that the maximum sensitivity has been produced by the optimal features instead of other sets of features.

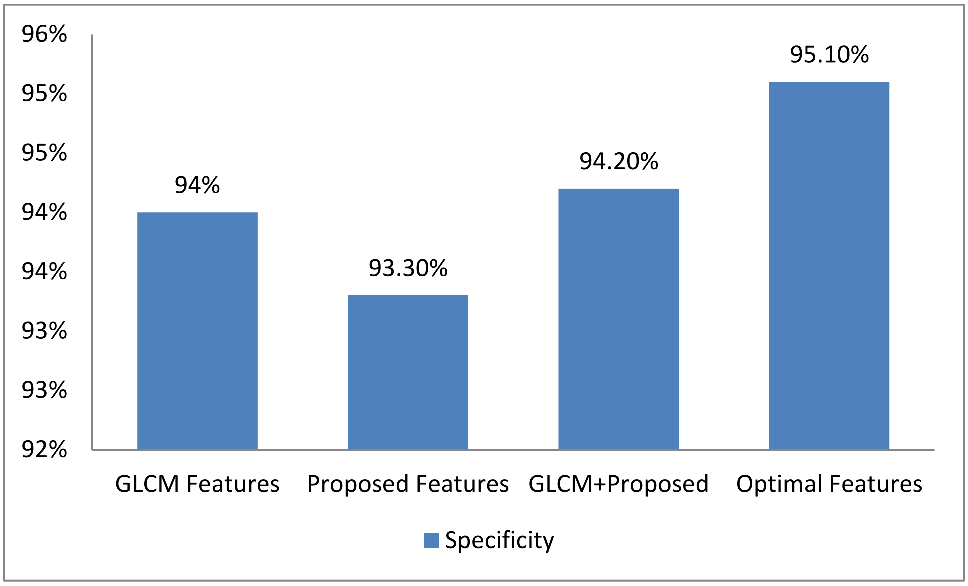

Figure 29 Comparison of specificity for different set of features.

Here, it can be observed that the maximum specificity has been produced by the optimal features instead of other sets of features.

Table 22 represents the classification results such as TP rate, FP rate, Precision, Recall, AUC etc. for the optimal features which gave the best accuracy of 94.5% at k-5 with KNN classifier.

Table 22 Classification results such as TP rate, FP rate, AUC, precision values.

Class	**TP Rate**	**FP Rate**	**Precision**	**Recall**	**F-Measure**
Dense	0.955	0.066	0.939	0.955	0.947
Fatty	0.934	0.045	0.952	0.934	0.943
Weighted Average	0.945	0.056	0.945	0.945	0.945

Chapter 7

Conclusion and Future Scope

7.1 CONCLUSION

This dissertation presents the classification of mammograms using texture analysis. 322 mammograms are classified into fatty, glandular and dense breast tissue type and 218 mammograms are classified into fatty and dense breast tissue type. In total, 16 features are extracted for texture analysis and out of which three are new gray-level intensity based features (white area, pixel count and contrast) that have been proposed in this work. Four sets of features are formed for classification. Classifiers used were naïve bayes, random forest and k-nearest neighbor. Conclusions of the work are listed below:

- All the three proposed features are endorsed by correlation feature selection (CFS) in the optimal features.
- When classifying the mammograms into three classes (fatty, glandular and dense), the set of proposed features give the best results having accuracy of 72.5 % with the classifier k-NN (k=5).
- When classifying the mammograms into fatty and dense, optimal features give the best results having accuracy of 94.5 % with the classifier k-NN (k=5).

It shows that texture features especially the proposed features can be effectively used for classification of mammograms. The classifiers used were best data mining algorithms; hence they are more reliable too. It also shows that dense breast tissue can be effectively classified from fatty tissue which is an indicator of having the chances of cancer.

7.2 FUTURE SCOPE

Even though it is proved that texture analysis is a very effective tool for classifying the mammograms into various classes, it does not differentiate completely. Some of the future scopes are listed below:

- Using different approaches like artificial intelligence, neural-networks, advanced computational techniques etc. the accuracy of classifying the mammograms into three classes can be improved.
- Glandular and dense type breast tissue look alike sometimes as the glandular type breast tissue is developing stage towards the dense type. So sometimes it is difficult to differentiate between them. Some new methods can be employed/developed to differentiate the two effectively.

References

[1] WHO Cancer Fact Sheets. (2009). [Online]. Available: http://www.who.int/mediacentre/factsheets/fs297/en/index.html

[2] M. Althius, J. Dozier, W. Anderson, S. Devesa, and L. Brinton "Global trends in breast cancer incidence and mortality 1973-1997," Int. J. Epidemiol, vol. 334, no. 2, pp. 405-412, 2005.

[3] E. Feurer, L. Wun, Devcan "Proability of developing or dying of cancer." Version 4.0, 1999.

[4] S. Obenauer, C. Sohns, C. Werner, E. Grabbe "Imapact of breast density on computer-aided detection in full-field digital mammography," Journal of Digital Imaging, Vol. 19, No. 3, pp. 258-263, 2006.

[5] N. Wolfe, "Risk for breast cancer development determined by mammographic parenchymal pattern," Cancer, vol. 37, pp. 2486–2492, 1976.

[6] B. Manjunath, W. Ma, "Texture feature for browsing and retrieval of image data", IEEE Trans PAMI, pp. 837-842, 1996.

[7] J. Juhl, J. Paul & Juhl's "Essentials of roentgen interpretation". 4th edition. Harper & row, pp. 340-345, 1982.

[8] L. Giger, F. Yin, and K. Doi., "Investigation of methods for the computerised detection and analysis of mammographic masses," SPIE Medical Imaging and Image Processing IV, 1233:183–184, 1990.

[9] D. Davies and D. Dance, "Automatic computer detection of clustered calcifications in digital mammograms," Phys. Med. Biol, 35:1111–1118, 1990.

[10] H. Chan, K. Doi, and C. Vyborny, "Improvements in radiologists' detection of clustered microcalcifications on mammograms: the potential of computeraided diagnosis. Investigative Radiology," 25:1102–1110, 1990.

[11] N. Karssemeijer, "A stochastic model for automated detection of calcifications in digital mammograms," Image and Vision Computing, 10:369–375, 1992.

[12] P. Miller and S. Astley, "Classification of breast tissue by texture analysis," Image and Vision Computing, 10:277–283, 1992.

[13] D. Wei, H. Chan, M. Helvie, B. Sahiner, N. Petrick, D. Adler, and M. Goodsitt, "Classification of mass and normal breast tissue on digital mammograms: multiresolution texture analysis," Medical Physics., vol. 22, pp. 1501-13, 1995.

[14] D. Wei, H. Chan, N. Petrick, B. Sahiner, M. Helvie, D. Adler, and M. Goodsitt, “False-positive reduction technique for detection of masses on digital mammograms: global and local multiresolution texture analysis,” Medical Physics., vol. 24, pp. 903-14, 1997.

[15] T. Ole, Gulsrud, and E. Loland, “Multichannel filtering for texture extraction in digital mammograms,” 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1996.

[16] J. Byng, N. Boyd, E. Fishell, R. Jong, and M. Yaffe, “Automated analysis of mammographic densities,” Phys. Med. Biol., vol. 41, pp. 909– 923, 1996.

[17] K. Bovis and S. Singh, “Detection of masses in mammograms using texture features,” 15th International Conference on Pattern Recognition (ICPR’00), 2:2267, 2000.

[18] G. Brake, N. Karssemeijer, and J. Hendriks, "An automatic method to discriminate malignant masses from normal tissue in digital mammograms," Physics in Medicine & Biology., vol. 45, pp. 2843-2857, 2000.

[19] A. Oliver, J. Freixenet, R. Marti, J. Pont, E. Perez, E. R. E. Denton, and R. Zwiggelaar, “A Novel Breast Tissue Density Classification Methodology,” IEEE Trans. on Information Technology in Biomedicine Vol. 12, No. 1, pp. 55-65 Jan. 2008.

[20] T. Docusse, A. Pereira, N. Marranghello, ,"Microcalcification border characterization," Engineering in Medicine and Biology Magazine, IEEE , vol.28, no.5, pp.41- 43, September-October 2009.

[21] M. Eltoukhy, I. Faye, B. Samir, "Curvelet based feature extraction method for breast cancer diagnosis in digital mammogram," Intelligent and Advanced Systems (ICIAS), 2010 International Conference on, vol., no., pp.1-5, 15-17 June 2010.

[22] N. Saidin, U. Ngah, H. Sakim, D. NikSiong, M. Hoe, "Density based breast segmentation for mammograms using graph cut techniques," TENCON 2009 - 2009 IEEE Region 10 Conference , vol., no., pp.1-5, 23-26 Jan. 2009.

[23] N. Saidin, U. Ngah, H. Sakim, D. NikSiong, M. Hoe, I. Shuaib, "Density Based Breast Segmentation for Mammograms Using Graph Cut and Seed Based Region Growing Techniques," Computer Research and Development, 2010 Second International Conference on , vol., no., pp.246-250, 7-10 May 2010.

[24] K. Bovis and S. Singh, "Classification of mammographic breast density using a combined classifier paradigm," in Proc. Med. Image Understanding Anal. Conf., pp. 177–180, 2002.

[25] R. Martí, R. Zwiggelaar, and C. Rubin, "A novel similarity measure to evaluate image correspondence," 15th International Conference on Pattern Recognition (ICPR'00), 3:3171, 2000.

[26] L. Blot and R. Zwiggelaar, "Extracting background texture in mammographic images: a co-occurrence matrix based approach," Proceedings of the 5th International Workshop on Digital Mammography, Toronto(Canada), pages 142–148, 2000.

[27] L. Blot, R. Zwiggelaar, and C. Boggis, "Enhancement of abnormal structures in mam- mographic images," Proceedings of Medical Image Understanding and Analysis, pages 125–128, 2000.

[28] L. Blot and R. Zwiggelaar, "Background texture extraction for the classification of mammographic parenchymal patterns," Medical Image Understanding and Analysis, pages 145–148, 2001.

[29] N. Youssry, F. Abou-Chadi, and A. El-Sayad, "Early detection of masses in digitized mammograms using texture features and neuro-fuzzy model," 4th Annual IEEE Conf. on Information Technology Applications in Biomedicine, 2003.

[30] J. Martí, J. Freixenet, X. Mu noz, and A. Oliver, "Active region segmentation of mam- mographic masses based on texture, contour and shape features," Springer-Verlag Berlin Heidelberg, LNCS 2652:478–485, 2003.

[31] M. Jirari, "A computer aided detection system for digital mammograms based on radial basis functions and feature extraction techniques," IEEE Engineering in Medicine and Biology 27th Annual Conference, 2005.

[32] L. Lena, P. Sakellaropoulos, M. Kristalli, S. Skiadopoulos, A. Karahaliou, I. Boniatis, and G.Panayiotakis, "Multiresolution feature analysis for differentiation of breast masses from normal tissue," 1st International Conference on Experiments/Process/System Modeling/Simulation/Optimization, 2005.

[33] M. Lyra, S. Lyra, B. Kostakis, S. Drosos, and C. Georgosopoulos, "Digital mammography texture analysis by computer assisted image processing," IEEE

International Workshop on Imaging Chania Greece September 2, pages 223–227, 2008.

[34] A. Karahaliou, I. Boniatis, P. Sakellaropoulos, S. Skiadopoulos, G. Panayiotakis, and L. Costaridou, “Can texture of tissue surrounding microcalcifications in mammography be used for breast cancer diagnosis,” Nuclear Instruments and Methods in Physics Research A 580, pages 1071–1074, 2007.

[35] A. Karahaliou, I. Boniatis, G. Skiadopoulos, F. Sakellaropoulos, N. Arikidis, E. A. Likaki, G. Panayiotakis, and L. Costaridou, “Breast cancer diagnosis: Analyzing texture of tis- sue surrounding microcalcifications,” IEEE Transactions on information technology in Biomedicine, 12:6, 2008.

[36] R. Panchal, B. Verma, "Classification of breast abnormalities in digital mammograms using image and BIRADS features in conjunction with neural network," Neural Networks, 2005. IJCNN '05. Proceedings. 2005 IEEE International Joint Conference on, vol.4, no., pp.2487-2492 vol. 4, July 31 2005-Aug. 4 2005.

[37] American Cancer Society, “Breast Cancer Facts and Figures 2011-2012,” Atlanta, Ga: American Cancer Society, 2011.

[38] N. Breen, J. Gentleman, J. Schiller, ‘Update on mammography trends: comparisons of rates in 2000, 2005, and 2008,” Cancer, 117(10):2209-2218, 2011.

[39] J. Lewin, C. D’Orsi, R. Hendrick, et al. “Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer,” Am J Roentgenol, 179:671-677, 2002.

[40] E. Pisano, C. Gatsonis, E. Hendrick, et al. “Diagnostic performance of digital versus film mammography for breast-cancer screening,” N Engl J Med., 353:1773-1783, 2005.

[41] American College of Radiology—Breast Imaging Reporting and Data system, 3rd edition. Reston, VA: American College of Radiology, 1992.

[42] N. Boyd, J. Byng, R.Long, E. Fishell, L.Little, A.Miller, G.Lockwood, D. Tritchler, M. Yaffe, “Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening study,” J Natl Cancer Inst 87:670-675, 1995.

[43] E. Tao, J. Sklansky, “Analysis of mammograms aided by database of images of calcifications and textures,” Proc SPIE 2710:988Y995, 1996.

[44] H. Alto, R. Rangayyan, J. Desautels, "Content-based retrieval and analysis of mammographic masses," J Electron Imaging 14(2):Article 023016:1Y17, 2005.

[45] R. Hubbard, K. Kerlikowske, C. Flowers, et al., "Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: a cohort study," Ann Intern Med, 155:481-492, 2011.

[46] K. Doi, H. MacMahon, M. Giger, K. Hoffmann, "Computer-aided diagnosis and its potential impact on diagnostic radiology," Computer-Aided Diagnosis in Medical imaging. Amsterdam, The Netherlands: Elsevier Science, pp 11-20, 1999.

[47] H. Tamura, S. Mori, and Y. Yamawaki, "Textural Features Corresponding to Visual Perception," IEEE Transactions on Systems, Man, and Cybernetics, SMC-8, pp. 460-473, 1978.

[48] J. Sklansky, "Image Segmentation and Feature Extraction," IEEE Transactions on Systems, Man and Cybernetics, SMC-8, pp. 237-247, 1978.

[49] T. Ojala and M. Pietikäinen, "Texture Classification. Machine Vision and Media Processing Unit," University of Oulu, Finland. Available at: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OJALA1/texclas.htm. January, 2004.

[50] R. Haralick, L. Watson, "A facet model for image data," Comput. Vision Graphics Image Prcoess, 15, 113-129, 1981.

[51] C. Gotlieb and H. Kreyszig, "Texture Descriptors based on Co-ocurrence Matrices", Computer Vision, Graphics, and Image Processing, Vol. 51, pp. 70-86, 1990.

[52] S. Peckinpaugh, "An Improved Method for Computing Gray-Level Coocurrence Matrix Based Texture Measures", Computer Vision, Graphics, and Image Processing; Graphical Models and Image Processing, Vol. 53, pp. 574-580, 1991.

[53] J. Suckling, J. Parker, D. R. Dance, S. M. Astley, I. Hutt, C.R.M. Boggis, I. Ricketts, E. Stamatakis, N. Cerneaz, S.L. Kok, P. Taylor, D. Betal and J. Savage. "The mammographic image society digital mammogram database," In Proc. International Workshop on Digital Mammography, pages 211–221, 1994.

[54] J. Gotlieb and H. Kreyzig, "Texture Descriptors based on Co-ocuurence Matrices," Computer Vision, Graphics and Image Processing, Vol. 51, pp. 70-86, 1990.

[55] D. He, L. Wang and J. Jiubert. "Texture Feature Extraction," Pattern Recognition Letters, Vol. 6, pp. 269—273, 1987.

[56] J. Gennari, P. Langely and D. Fisher, “Models of incremental concept formation,” Artifical Intelligence, (40):11-61, 1989.

[57] R. Kohavi and G. John, “Wrappers for feature subset selection,” Artificial Intelligence, special issue on relevance, 97(1-2):273-324, 1996.

[58] M. Hall, “Correlation based Feature Selection for Machine Learning”, Department of Computer Science, University of Waikato, 1999.

[59] T. Cover, P. Hart, "Nearest neighbor pattern classification," Information Theory, IEEE Transactions on , vol.13, no.1, pp.21-27, January 1967.

[60] T. Ho, “Random decision Forest,” Document Analysis and Recognition, 1995, Proceedings of third International Conference on, vol. 1, pp. 278-282, 14-16 Aug 1995.

[61] A. Prasad, L. Iverson, A. Liaw, “Newer classification and regression techniques: bagging and random forests for ecological prediction,” Ecosystems 9(2):181-199, 2006.

[62] P. Langley and S. Sage, “Induction of selective Bayesian classifiers,” In Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence, Seattle, W.A, 1994. Morgan Kaufmann.