CHAPTER 3

IMAGE SEGMENTATION USING INTUITIONISTIC FUZZY C MEANS CLUSTERING WITH SPATIAL INFORMATION

This chapter introduces the concept of FCM, IFCM and fuzzy entropy and proposes a segmentation algorithm based on IFCM by incorporating the spatial information. It also details the segmentation results of the proposed system with the existing cluster based methods.

3.1 INTRODUCTION

CAD systems for mammogram involve the detection of suspicious region or ROI, feature extraction, feature selection and classification steps. Accurate segmentation of ROI allows the precise extraction of lesion features which improves the performance of CAD. The segmentation techniques can be divided into supervised and unsupervised techniques. Supervised segmentation relies on prior knowledge about the object and background regions to be segmented. This prior information is used to determine if specific regions are present within an image or not.

Alternatively, unsupervised segmentation partitions an image into a set of regions which are distinct and uniform with respect to specific properties such as gray level, texture or color. Classical approaches to solving unsupervised segmentation are divvied into three major groups: 1) region based segmentation, which divides the image into homogeneous and spatially connected regions, 2) contour based segmentation, which relies on the
boundaries of the regions and 3) clustering methods, which groups pixels of the same properties and results in non-connected regions. The most commonly used clustering algorithms are K-means, FCM and IFCM (Martins et al 2009; Pham et al 1999).

K-means is the simplest clustering algorithm for partitioning a dataset into K subsets (Martins et al 2009). The traditional hard clustering method restricts each point of the dataset exclusively to one cluster. As a consequence, the segmentation results are often very crisp; each pixel of the image belongs to exactly just one class. Moreover, the algorithm is significantly sensitive to the initial partition and is more likely to converge to a local minimum resulting from improper initialization.

The FCM incorporates Fuzzy Set (FS) theory into the K-means clustering, which allows one data point to belong to two or more clusters. When segmenting, the clustering may be fuzzy or crisp but the images are considered to be fuzzy due to the uncertainty present in terms of ambiguity in class definitions, regions or boundaries and imprecise gray levels present in them. So fuzzy clustering is the most widely used method where an element may have partial membership grades in more than one fuzzy cluster. The main problem of this approach lies on the expert’s choice of the membership functions that assign each pixel either to the background or to the object. Moreover, this choice has proven to be of utmost importance in terms of the performance of the algorithm.

To overcome the in FCM, IFCM clustering incorporates intuitionistic fuzzy index values to represent the hesitance of an expert on determining whether a pixel of the image belongs to which cluster. In FS, the non-member degree is the complement of the membership degree, whereas in
Intuitionistic Fuzzy Set (IFS), the non-membership degree is less than or equal to the complement of the membership degree due to the hesitation degree (Chaira 2011). Since it outperforms the other clustering techniques, IFCM clustering with spatial information is used for segmenting the suspected region.

3.2 SEGMENTATION USING FUZZY C MEANS CLUSTERING

FCM is a method of clustering in which one piece of data belongs to two or more clusters. Let \( X = (x_1, x_2, \ldots, x_N) \) be denoted as an image with \( N \) pixels to be partitioned into \( c \) clusters, where \( x_i \) represents multispectral features. This FCM clustering algorithm is an iterative function which aims at minimizing the intercluster similarity and minimizing the cost function. The objective function or the cost function of the FCM clustering is defined as

\[
J(U, V) = \sum_{j=1}^{N} \sum_{i=1}^{c} (\mu_{ij})^m \|x_j - v_i\|^2
\]  

(3.1)

where \( \mu_{ij} \) represents the membership value of pixel \( x_j \) in the \( i \)th cluster; \( v_i \) is the \( i \)th cluster centre, \( \| \cdot \| \) is a norm metric and \( m \) is a constant. The parameter \( m \) controls the fuzziness of the resulting partition, and the value of \( m \) is 2 in this study. The membership \( u_{ij} \) and cluster centers \( v_i \) are updated according to the Equations (3.2) and (3.3) respectively.

\[
u_{ij} = \frac{1}{\sum_{k=1}^{c} \left( \frac{\|x_j - v_i\|}{\|x_j - v_k\|} \right)^{2/(m-1)}}
\]  

(3.2)

\[
v_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m}
\]  

(3.3)
The membership function represents the probability that a pixel belongs to a specific cluster. The sequence of steps used to segment the image is explained in Algorithm 3.1.

**Algorithm 3.1: steps for FCM clustering**

Step 1: Randomly select c cluster centers.

Step 2: Calculate the fuzzy membership using Equation (3.2)

Step 3: Compute the cluster centre using Equation (3.3)

Step 4: Update the objective function or cost function using Equation (3.1)

Step 4: Repeat step2 and step4 until minimum cost is achieved.

**3.3 PROPOSED METHODOLOGY**

A new method used for segmenting and detecting the breast cancer in mammogram images is presented in Figure 3.1.

**3.3.1 Preprocessing**

In a typical mammogram, several areas such as image background, digitization noises, informative marks and pectoral region and so on are present. Prior to segmentation, all the noises should be removed (Nagi et al 2010). In order to remove the noises, the dilation and erosion operations are used. Dilation and erosion are basic morphological operations defined in Equations (3.4) and (3.5) respectively.

\[
A \oplus B = \left\{ s \left| \left[ (B)_z \cap A \right] \subseteq A \right\} \right. 
\]  \hspace{1cm} \text{(3.4)}

\[
A \ominus B = \{ s \left| (B)_z \subseteq A \right\} \right. 
\]  \hspace{1cm} \text{(3.5)}
First, the image is converted into black and white. Then, a sequence of morphological operations, dilation and erosion, are applied for removing the label and the background. Similarly, the pectoral muscle region is removed to increase the detection performance. A window of size $7 \times 7$ pixels...
is extracted for each pixel centered at the pixel location. The median value is computed for this window. The intensity value of the center pixel is replaced with the median value. To remove the pectoral muscle region, initially the histogram is generated for the mammogram image. The global optimum in the histogram is selected as the threshold value. When the MLO view is correctly imaged, the pectoral region should always appear as a high-intensity and triangular region across the upper posterior margin of the image. In many cases the upper part of the boundary is a sharp intensity edge while the lower part is more likely to be a texture edge due to the fact that it is overlapped by fibroglandular tissue. After finding the global optimum value the image of the left breast is scanned from top left to right or top right to left in the case of right breast in the triangular region across the upper posterior margin of the mammogram image. The intensity values greater than this threshold are changed into black and the gray values smaller than the threshold are maintained as such to convert the pictorial region into black region. The various stages of image preprocessing are shown in Figure 3.2.

![Figure 3.2 Various stages of preprocessing](image)

(a) Original image (b) Image after removing artifacts and background (c) Image after removing pectoral region
### 3.3.2 Segmentation using Intuitionistic Fuzzy C Means Clustering with Spatial Information (sIFCM)

#### Intuitionistic Fuzzy Set

sIFCM clustering algorithm is based on IFS theory. IFS is the extension of the FS which generates only membership function $\mu(x), x \in X$, whereas intuitionistic fuzzy set defined by Atanassov (1999) considers both membership $\mu(x)$, and non membership $\nu(x)$. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a set of elements; an IFS $A$ in $X$ is given by an ordered triple as

$$A = \{(x, \mu_A(x), \nu_A(x))|x \in X\}$$  \hspace{1cm} (3.6)

where $\mu_A, \nu_A: X \rightarrow [0,1]$ should satisfy the condition

$$0 \leq \mu_A(x) + \nu_A(x) \leq 1$$  \hspace{1cm} (3.7)

for all $x \in X$. The values of $\mu_A(x)$ and $\nu_A(x)$ denote the degree of membership and non membership of element $x \in X$ to $A$, respectively. The membership value indicates the degree of belongingness of that element to the set, whereas the non-membership value indicates the degree to which that element does not belong to the set.

For all IFS, Atanassov (1999) also indicated a hesitation margin (or intuitionistic fuzzy index), $\pi_A(x)$, which arises due to the lack of knowledge in defining the membership degree of each element $x$ in the set $A$ and is given by:

$$\pi_A(x) = 1 - \mu_A(x) - \nu_A(x)$$  \hspace{1cm} (3.8)
It is obvious that $0 \leq \pi_A(x) \leq 1$, for all $x \in X$.

Intuitionistic fuzzy generator or fuzzy complement is created from Yager generating function (Chaira 2011). The fuzzy complement is defined as

$$N(\mu(x)) = g^{-1}(g(1) - g(\mu(x)))$$

(3.9)

where $g(.)$ is an increasing function and $g:[0,1] \rightarrow [0,1]$. Yager class can be generated by substituting the Equation (3.10) with the Equation (3.9).

$$g(x) = x^\alpha$$

(3.10)

So Yager’s intuitionistic fuzzy complement is written as

$$N(x) = (1 - x^\alpha)^{1/\alpha},$$

(3.11)

$\alpha > 0$, where $N(1) = 0, \ N(0) = 1$

Non-membership values are calculated from Yager’s intuitionistic fuzzy complement $N(x)$. Thus, with the help of Yager’s intuitionistic fuzzy complement, IFS and hesitation degree are defined as

$$A^\text{IFS}_A = \{x, \mu_A(x), (1 - \mu_A(x)^\alpha)^{1/\alpha} | x \in X\}$$

(3.12)

$$\pi_A(x) = 1 - \mu_A(x) - (1 - \mu_A(x)^\alpha)^{1/\alpha}$$

(3.13)

In this study the value of $\alpha$ is selected as 0.80 empirically.
**Intuitionistic Fuzzy Image**

The basic requirement in intuitionistic fuzzy image processing is the computation of membership and non-membership degrees. Let X be an image of $M \times N$ pixels having L gray levels ranging between 0 and L-1. Then the image is initially fuzzified using the Equation (3.14).

$$\mu_A(p_{ij}) = \frac{p_{ij} - p_{min}}{p_{max} - p_{min}}$$

(3.14)

where $p_{ij}$ is the pixel value at (i,j)th position ranging from 0 to L-1; i=0,1,2,… N-1 and j=0,1,2,…M-1; $p_{min}$ and $p_{max}$ are the minimum and maximum gray values of the image X.

The degree of non-membership and hesitation value is calculated using Equations (3.15) and (3.16) respectively.

$$\nu_A(p_{ij}) = (1 - \mu_A(p_{ij})^\alpha)^{1/\alpha}$$

(3.15)

$$\pi_A(p_{ij}) = 1 - \mu_A(p_{ij}) - (1 - \mu_A(p_{ij})^\alpha)^{1/\alpha}$$

(3.16)

Each image pixel is associated with three numerical values as defined below.

- **Membership $\mu_A(x)$**: A value that represents the membership obtained by means of membership function associated with the set that indicates the expert’s knowledge of the image.

- **Non-membership $\nu_A(x)$**: A value that represents the non-membership obtained by means of non-membership function associated with the set that indicates the ignorance of the expert’s decision.
Hesitation measure $\pi_A(x)$: A value that represents the hesitation measure and is obtained by using Equation (3.16).

Finally, the image in intuitionistic fuzzy is defined as

$$A_{IFS} = \{p_{ij}, \mu_A(p_{ij}), v_A(p_{ij})\}, \quad p_{ij} \in (0,1,...L - 1)$$

(3.17)

where $p_{ij}$ is the pixel value at (i,j)th position, i=0,1,2,… N-1 and j=0,1,2,..M-1.

In addition to the pixel value, the image has features like pixel mean, standard deviation and entropy. A size of 3x3 square windows is considered to calculate the mean, standard deviation and entropy.

**Intuitionistic Fuzzy Entropy**

Uncertainty is found to exist in real medical images. Intuitionistic Fuzzy Entropy (IFE) defines the amount of vagueness or fuzziness in FS. Many researchers have defined IFE in different ways. Burillo & Bustince (1996) have first proposed the entropy of IFS. A real function $IFE: IFS(X) \rightarrow \mathbb{R}^+$ is entropy on $IFS(X)$ if it satisfies the following properties.

1. $IFE(A)=0$ if $1 - [\mu_A(x_i) + v_A(x_i)] = 0$

   which implies $\mu_A(x_i) + v_A(x_i) = 1$, for all $x_i$ and thus $A \in FS$.

2. $IFE(A)=n$ if $1 - [\mu_A(x_i) + v_A(x_i)] = 1$

   This occurs iff $\mu_A(x_i) = v_A(x_i) = 0$ for all $x_i$.

3. $IFE(A)=IFA(A^c)$ for all $A \in IFS(X)$.

4. $IFE(A) \geq IFE(B)$ if $\mu_A(x_i) \leq \mu_B(x_i)$ and $v_A(x_i) \leq v_B(x_i)$ for all $x_i$.

Thus $\mu_A(x_i) + v_A(x_i) \leq \mu_B(x_i) + v_B(x_i) \Rightarrow \pi_A(x_i) \geq \pi_B(x_i)$. 
Pal & Pal (1991) have analyzed the classical Shannon information theory and introduced entropy. For a probability distribution, \( p = p_1, p_2, \ldots, p_n \), the exponential entropy is defined as

\[
H = \sum_{i=1}^{n} p_i e^{1-p_i} \quad (3.18)
\]

for intuitionistic fuzzy cases, if \( \mu_A(x), \nu_A(x), \pi_A(x) \) are the membership, non-membership and hesitation degrees of the elements of the set \( X \), then IFE, that denotes the degree of intuitionism in FS is defined as

\[
IFE(A) = \sum_{i=1}^{n} \pi_A(x_i)e^{1-\pi_A(x_i)} \quad (3.19)
\]

IFE(A) satisfies the conditions of IFE (Burillo & Bustince 1996).

**Intuitionistic Fuzzy C-Means Clustering**

IFCM algorithm assigns pixels to each category by using membership, non-membership and hesitancy values. Let \( X = (x_1, x_2, \ldots, x_{M \times N}) \) be an image with \( N \times M \) pixels to be partitioned into \( c \) clusters where \( x_i \) represents multispectral (features) data. The IFCM clustering algorithm is an iterative function and aimed at minimizing the inter cluster similarity and IFE. The objective function of IFCM clustering is defined as follows.

\[
J = \sum_{i=1}^{c} \sum_{k=1}^{n} u_{ik}^m d(x_k, v_i)^2 + \sum_{i=1}^{c} \pi_i^* e^{1-\pi_i^*} \quad \text{with} \quad m = 2 \quad (3.20)
\]
where $c$ is the number of clusters, $n$ is the number of data points, $u_{ik}^*$ is the intuitionistic fuzzy membership matrix, $v_i$ is the cluster center; $d(x_k, v_i)$ is the distance measure between data points and cluster centre; and $\pi_i^* e^{1-\pi_i^*}$ is the fuzzy entropy.

\begin{equation}
    u_{ik}^* = \frac{1}{\sum_{j=1}^{c} [d_{ik}^2 / d_{jk}^2]^{1/m-1}} + \pi_{ik} 
\end{equation}

\begin{equation}
    v_i^* = \frac{\sum_{k=1}^{n} u_{ik}^* x_{ik}}{\sum_{k=1}^{n} u_{ik}^*} 
\end{equation}

Normalized Euclidian distance measure is used to calculate the distance between the data points (A) and cluster centre (B) and defined as

\begin{equation}
    qIFS(A, B) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\mu_A(x_i) - \mu_B(x_i))^2 + (v_A(x_i) - v_B(x_i))^2} 
\end{equation}

During implementation, matrix $v^*$ is randomly initialized, and then $u^*$ and $v^*$ are updated through an iterative process using Equations (3.21) and (3.22) respectively.

**IFCM Clustering with Spatial Information (sIFCM)**

Generally, the IFCM algorithm does not fully utilize the spatial information in the image. In the proposed research, an IFCM algorithm that incorporates spatial information into the membership function for clustering is proposed. The spatial function is the summation of the membership function in the neighborhood of each pixel under consideration. The advantages of the new method are: (1) yielding regions more homogeneous than those of other methods, (2) reducing the spurious blobs, (3) removing noisy spots, and (4) being less sensitive to noise than the other techniques.
This sIFCM technique is a powerful method for noisy image segmentation and works for both single and multiple-feature data with spatial information. One of the important characteristics of an image is that the neighboring pixels are highly correlated. This spatial relationship is important in clustering and to exploit the spatial information, a spatial function is defined as

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik}^* \quad (3.24)$$

where $NB(x_j)$ represents a square window centered on pixel $x_j$ in the spatial domain and $h_{ij}$ represents the probability that a pixel $x_j$ belongs to the $i$th cluster. The spatial function of a pixel for a cluster is large if the majority of its neighborhood belongs to the same cluster. By incorporating spatial information, the new membership function is defined as follows

$$u_{ij}^* = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^c u_{ik}^p h_{ij}^q} \quad (3.25)$$

where $p$ and $q$ are the parameters to control the relative importance of both functions. The value of $p$ and $q$ are selected empirically as 0.85 and 1.5 respectively. Thus the modified objective function of sIFCM is defined as follows.

$$J_{sIFCM} = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^{*m} d(x_k, v_i)^2 + \sum_{i=1}^c \pi_i^* e^{1-\pi_i^*} \quad \text{with } m = 2 \quad (3.26)$$
Algorithm 3.2: Steps for sIFCM clustering

Step 1 : Determine initial centroids by selecting \( c \) random intuitionistic fuzzy objects; initialize the membership matrix \( u_{ij} \)

Step 2 : Update the membership matrix using Equation (3.21)

Step 3 : Update the hesitation matrix

Step 4 : Update the spatial relationship using the Equation (3.24)

Step 5 : Update the membership matrix by incorporating spatial relationship using Equation (3.25)

Step 6 : Update the intuitionistic fuzzy entropy using Equation (3.19)

Step 7 : Update the cluster centre

Step 8 : Find the new value of the objective function using Equation (3.25).

Step 9 : Repeat step 3 to step 8 till convergence

3.3.3 Feature Extraction

Application of data mining techniques to image data involves a number of challenges related to the representation of images into an appropriate format which permits for mining knowledge from the images using data mining techniques. The feature is defined as a function of one or more measurements, each of which specifies some quantifiable property of an object, and is computed such that it quantifies some significant characteristics of the object. For each ROI, the following features are extracted and a feature database is constructed for each ROI of the mammogram.
First Order Statistics Texture Features

First order statistical texture features provides different statistical properties of an image. Intensity histogram features or first order statistics like mean, standard deviation, skewness, entropy and kurtosis are extracted from each ROI of the images (Karahaliou et al 2008).

Discrete Wavelet Transformation (DWT)

DWT is a powerful mathematical tool and one of the multiscale techniques for image analysis. By applying DWT, the ROI is divided into 4 subbands namely, Low Low (LL), Low High (LH), High Low (HL) and High High (HH) components. Two level DWT is applied in the present study. Figure 3.3 illustrates the two level DWT and the iteration of only LL subband for further level decomposition.

<table>
<thead>
<tr>
<th>LL1</th>
<th>LH1</th>
<th></th>
<th>LL2</th>
<th>LH2</th>
<th>LH1</th>
</tr>
</thead>
<tbody>
<tr>
<td>HL1</td>
<td>HH1</td>
<td></td>
<td>HL2</td>
<td>HH2</td>
<td>HH1</td>
</tr>
</tbody>
</table>

Figure 3.3 Discrete wavelet decomposition

Second Order Statistical Texture features

The GLCM is a well known statistical tool for extracting second order texture information from images (Haralick et al 1973). It involves two steps to generate spatial features: extraction of spatial information and computation of statistical features.
At first, the spatial information of each subband is calculated on a pixel neighbourhood defined by a moving window of a given size. GLCM is computed on the basis of joint conditional probability density functions $P(i, j|d, \theta)$. The four orientations that are mainly focused during the generation of the matrix are $0^\circ$ or horizontal direction, $45^\circ$ direction, $90^\circ$ or vertical direction and $135^\circ$ direction as shown in Figure 3.4.

![Figure 3.4. Four orientations towards the generation of GLCM](image)

The function $p(i, j|d, \theta)$ is the probability that two pixels, which are located with an inter sample distance $d$ and a direction $\theta$, have a gray level $i$ and a gray level $j$. The estimated joint conditional probability density functions are defined as in Equations (3.27) to (3.30).

\[
p(i, j|d, 0^\circ) = \frac{\#\{((k, l), (m, n)) \in (L_x \times L_y) \times (L_x \times L_y): k = m |l - n| = d, l(k, l) = i, l(m, n) = j\}}{T(d, 0^\circ)} \tag{3.27}
\]

\[
p(i, j|d, 45^\circ) = \frac{\#\{((k, l), (m, n)) \in (L_x \times L_y) \times (L_x \times L_y): (k - m = d, l - n = -d)\text{ or } (k - m = -d, l - n = d), l(k, l) = i, l(m, n) = j\}}{T(d, 45^\circ)} \tag{3.28}
\]
where \( \# \) denotes the number of elements in the set, \( I(x, y) \) is the image intensity at the point \((x, y)\) and \( T(d, \theta) \) stands for the total number of pixel pairs within the image which have the \( d \) inter sample distance and direction \( \theta \).

Each of the estimated joint probability density functions can be written in matrix form; the gray level cooccurrence matrix, \( \Phi(d, \theta) \), as defined in Equation (3.31).

\[
\Phi(d, \theta) = [p(i, j|d, \theta)], \quad 0 \leq i, j \leq N_g
\]

where \( N_g \) is the maximum gray level. If a texture is coarse and \( d \) is small compared to the sizes of the texture elements, the pairs of points at the inter sample distance \( d \) should usually have similar gray level. This means that the probability distribution in the matrix \( \Phi(d, \theta) \) is concentrated on or near its diagonal. On the other hand, for a fine texture, the gray levels of the points separated by the distance \( d \) should be quite different so that the probability distribution in \( \Phi(d, \theta) \) is distributed away from its diagonal.
The GLCM matrix is calculated for three high frequency subbands and for the directions of 0°, 45°, 90° and 135°, with the distances 1, 2, 3, and 4.

The second step is to compute thirteen statistical features such as angular second moment or energy, contrast or inertia, correlation, sum of square or variance, inverse different moment or homogeneity, sum average, sum variance, sum entropy, entropy, difference entropy, difference variance, information measure of correlation 1 and information measure of correlation 2 from the GLCM (Mencattini 2010).

**Wavelet Energy Features**

Energy features are a measure of the dispersion of the wavelet coefficients, and strongly correlated. The wavelet energy signature reflects the distribution of energy along the frequency axis over scale and orientation and has proven to be very useful for gray-level texture characterization. The energy features extracted for both the levels of approximate and detailed components.

**3.3.4 Classification using C4.5 Decision Tree**

A decision tree is generally constructed in a top-down manner. This construction begins at the root node where each attribute is evaluated using a statistical test to determine how well it can classify the training samples. The best attribute is chosen as the test at the root node of the tree. A descendant of the root node is then created for either each possible value of this attribute if it is a discrete valued attribute or each possible discretised interval of this attribute if it is a continuous valued attribute. Then the training samples are sorted to the appropriate descendant node. This process is repeated by the use of the training samples associated with each descendant node to select the
best attribute for testing at that point in the tree. This forms a greedy search for a decision tree in which the algorithm never backtracks to reconsider earlier node choices. Although it is possible to add a new node to the tree until all the samples that are assigned to one node belong to the same class, the tree is not allowed to grow to its maximum depth. A node is introduced to the tree only when there are a sufficient number of samples left. After constructing the complete tree, a tree pruning is carried out to avoid data over-fitting (Ruggieri 2001).

C4.5 builds decision trees from a set of training data using the concept of information entropy. The training data is a set with $t$ samples $(S = s_1, s_2, \ldots, s_t)$ of already classified samples. Each sample $s_i = x_1, x_2, \ldots, x_n$ is a vector where $x_1, x_2, \ldots, x_n$ represent attributes or features of the sample. The training data is augmented with a vector $C = c_1, c_2, \ldots, c_m$ where $c_1, c_2, \ldots, c_m$ represent the class to which each sample belongs.

At each node of the tree, C4.5 chooses one attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. Its criterion is the normalized information gain (difference in entropy) that results from choosing an attribute for splitting the data. The attribute with the highest normalized information gain is chosen to make the decision. The information gain ratio of an attribute $A$ relative to the sample set $S$ is defined as

$$\text{GainRatio}(A,S) = \frac{\text{Gain}(A,S)}{\text{SplitInformation}(A,S)} \quad (3.32)$$
\[ Gain(A, S) = Ent(S) - \sum_{a \in A} \frac{|S_a|}{|S|} \cdot Ent(S_a) \quad (3.33) \]

\[ SplitInformation(A, S) = -\sum_{a \in A} \frac{|S_a|}{|S|} \cdot \log_2 \frac{|S_a|}{|S|} \quad (3.34) \]

\[ Ent(S) = -\sum_{i \in C} \frac{|S_i|}{|S|} \cdot \log_2 \frac{|S_i|}{|S|} \quad (3.35) \]

where \( S_a \) is the subset of \( S \) for which the attribute \( A \) has the value \( a \) and \( C \) represents the classes in \( S \) (Witten et al 2011).

### 3.4 EXPERIMENTAL ANALYSIS

#### 3.4.1 Experimental Dataset

The images from the public and well known datasets like MIAS (Suckling et al 1994) and DDSM (Heath et al 2001) are used in this research.

The MIAS dataset consists of 322 mammogram images from 161 women. Bilateral, left and right MLO view mammograms are taken for each woman. The dataset belong to three big categories: normal, benign and malign which indicate different classes of abnormalities such as calcification (CALC), well-defined circumscribed masses (CIRC), speculated masses (SPIC), ill-defined masses (MISC), architectural distortion (ARCH), asymmetry (ASYM) and normal. It also includes the locations of abnormalities that may be present. Out of 322 images, 208 images are normal, 63 images are benign and 51 images are malignant. The different classes of abnormalities with respect to benign and malignant images are shown in Figure 3.5. In this chapter, 98 abnormal images (excluding asymmetry) with 103 TP ROIs and 100 normal images are considered.
Figure 3.5 Distribution of abnormalities in benign and malignant categories from MIAS dataset

The DDSM database contains approximately 2,500 studies. Each study provides four mammograms, comprising left and right MLO and left and right CC views, for most women. For the current study, 464 MLO view images which include 200 normal, 109 benign and 155 malignant images from the database are considered. The distribution with respect to benign and malignant used in this study are shown Figure 3.6.

Figure 3.6 Distribution of abnormalities in benign and malignant categories from DDSM dataset
3.4.2 Performance Evaluation

The performance of the proposed CAD system is measured based on the percentage of diagnostic decisions proven to be correct. However, this parameter strongly depends on the prevalence disease and it does not indicate FP error rates and FN error rates. To overcome these limitations, sensitivity, specificity and FP per image indices are also adopted. Sensitivity, also called recall rate, measures the proportion of actual positives which are correctly identified as such; the percentage of sick people who are correctly identified as having the condition. Specificity measures the proportion of negatives which are correctly identified; the percentage of healthy people who are correctly identified as not having the condition.

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + TP + FN} \tag{3.36}
\]

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \tag{3.37}
\]

\[
\text{Specificity} = \frac{TN}{TN + FP} \tag{3.38}
\]

\[
\text{FP per image} = \frac{FP}{\text{Total number of images}} \tag{3.39}
\]

where

TP => True Positive: a patient predicted with cancer when the patient actually has cancer.

TN => True Negative: a patient predicted healthy when the patient is actually healthy.

FN => False Negative: a patient predicted healthy when the patient actually has cancer

FP => False Positive: a patient predicted with cancer when the patient is actually healthy.
3.4.3 Results and Discussion

The Figure 3.7 shows the ROI extracted automatically using the FCM clustering, IFCM clusters and sIFCM clustering with ground truth information for some sample cases from and MIAS and DDSM dataset.

![Figure 3.7](image)

**Figure 3.7 Segmentation using ground truth information and using (a) FCM (b) IFCM (c) sIFCM**

Figure 3.7 shows the segmentation results in which white contour shows the segmentation result using ground truth information and red contour represents the segmentation result using sIFCM, FCM and IFCM. The segmentation results obtained by using ground truth information and using sIFCM are nearer when compared to FCM and IFCM.

Tables 3.1, 3.2 and 3.3 summarize the segmentation results of FCM, IFCM and the proposed sIFCM techniques for MIAS dataset.
### Table 3.1 Segmentation results using FCM method for MIAS dataset

<table>
<thead>
<tr>
<th></th>
<th>No. of images</th>
<th>No. of detected ROIs</th>
<th>No. of TP ROIs</th>
<th>No. of FP ROIs</th>
<th>No. of FN ROIs</th>
<th>No. of TN ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer Image</td>
<td>98</td>
<td>180</td>
<td>103</td>
<td>75</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Normal Image</td>
<td>100</td>
<td>120</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>120</td>
</tr>
<tr>
<td>Total</td>
<td>198</td>
<td>300</td>
<td>103</td>
<td>75</td>
<td>4</td>
<td>120</td>
</tr>
</tbody>
</table>

### Table 3.2 Segmentation results using IFCM method for MIAS dataset

<table>
<thead>
<tr>
<th></th>
<th>No. of images</th>
<th>No. of detected ROIs</th>
<th>No. of TP ROIs</th>
<th>No. of FP ROIs</th>
<th>No. of FN ROIs</th>
<th>No. of TN ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer Image</td>
<td>98</td>
<td>172</td>
<td>105</td>
<td>67</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Normal Image</td>
<td>100</td>
<td>105</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>105</td>
</tr>
<tr>
<td>Total</td>
<td>198</td>
<td>277</td>
<td>105</td>
<td>67</td>
<td>2</td>
<td>105</td>
</tr>
</tbody>
</table>

### Table 3.3 Segmentation results using the proposed sIFCM method for MIAS dataset

<table>
<thead>
<tr>
<th></th>
<th>No. of images</th>
<th>No. of detected ROIs</th>
<th>No. of TP ROIs</th>
<th>No. of FP ROIs</th>
<th>No. of FN ROIs</th>
<th>No. of TN ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer Image</td>
<td>98</td>
<td>160</td>
<td>107</td>
<td>53</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Normal Image</td>
<td>100</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>198</td>
<td>260</td>
<td>107</td>
<td>53</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
From Tables 3.1 to 3.3 shows that the segmentation methods FCM, IFCM and proposed sIFCM are found to register a total of 75, 67 and 53 FP ROIs, which is a rate of 0.77, 0.68 and 0.54 FP per image respectively. FCM and IFCM registered a total of 4 and 2 FNs whereas in sIFCM there is no FN ROI is registered.

![Graphical representation of the results of segmentation methods for MIAS dataset](image)

**Figure 3.8** Graphical representation of the results of segmentation methods for MIAS dataset

From Figure 3.8, it is observed that sIFCM reduces the FPs and FNs when compared to FCM and IFCM. Table 3.4, Table 3.5 and Table 3.6 summarize the segmentation results of FCM, IFCM and the proposed sIFCM techniques for DDSM dataset.

**Table 3.4 Segmentation results using FCM method for DDSM dataset**

<table>
<thead>
<tr>
<th>Class</th>
<th>No. of images</th>
<th>No. of detected ROIs</th>
<th>No. of TP ROIs</th>
<th>No. of FP ROIs</th>
<th>No. of FN ROIs</th>
<th>No. of TN ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer Image</td>
<td>264</td>
<td>420</td>
<td>264</td>
<td>156</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Normal Image</td>
<td>200</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>Total</td>
<td>464</td>
<td>620</td>
<td>264</td>
<td>156</td>
<td>0</td>
<td>200</td>
</tr>
</tbody>
</table>
Table 3.5 Segmentation results using IFCM method for DDSM dataset

<table>
<thead>
<tr>
<th></th>
<th>No. of images</th>
<th>No. of detected ROIs</th>
<th>No. of TP ROIs</th>
<th>No. of FP ROIs</th>
<th>No. of FN ROIs</th>
<th>No. of TN ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer Image</td>
<td>264</td>
<td>420</td>
<td>264</td>
<td>156</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Normal Image</td>
<td>200</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>Total</td>
<td>464</td>
<td>620</td>
<td>264</td>
<td>156</td>
<td>0</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 3.6 Segmentation results using the proposed sIFCM method for DDSM dataset

<table>
<thead>
<tr>
<th></th>
<th>No. of images</th>
<th>No. of detected ROIs</th>
<th>No. of TP ROIs</th>
<th>No. of FP ROIs</th>
<th>No. of FN ROIs</th>
<th>No. of TN ROIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer Image</td>
<td>264</td>
<td>412</td>
<td>264</td>
<td>148</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Normal Image</td>
<td>200</td>
<td>200</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>Total</td>
<td>464</td>
<td>612</td>
<td>264</td>
<td>148</td>
<td>0</td>
<td>200</td>
</tr>
</tbody>
</table>

From Tables 3.4, 3.5 and 3.6, it is found that for DDSM dataset FCM, IFCM and the proposed sIFCM segmentation methods register a total of 156 FP ROIs, 156 FP ROIs and 148 FP ROIs, which is a rate of 0.59, 0.59 and 0.56 FP per image respectively. The results are graphically represented in the following figure for MIAS dataset.
From Figure 3.9 it is observed that sIFCM reduces the FPs when compared to FCM and IFCM.

From Figure 3.7, it is clear that the morphology, size, shape and so on, of the segmented TP region with FCM, IFCM, sIFCM and using the ground truth information are not similar for all the cases. In order to evaluate the effectiveness of the segmentation results, various features like histogram features, second order statistical features and wavelet energy features are extracted from each detected TP ROIs and TN ROIs. Subsequently, the features are used to classify the ROI into normal, benign and malignant. The classifier C4.5 with 10 fold cross validation is used. The classification performance is shown in Table 3.7 and Table 3.8 for MIAS and DDSM dataset respectively.
Table 3.7 Classification performance of segmented ROIs from MIAS dataset

<table>
<thead>
<tr>
<th>Segmentation Method</th>
<th>Performance Measures</th>
<th>FCM</th>
<th>IFCM</th>
<th>sIFCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (%)</td>
<td></td>
<td>81</td>
<td>83</td>
<td>89</td>
</tr>
<tr>
<td>Sensitivity</td>
<td></td>
<td>0.77</td>
<td>0.76</td>
<td>0.86</td>
</tr>
<tr>
<td>Specificity</td>
<td></td>
<td>0.85</td>
<td>0.9</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table 3.8 Classification performance of segmented ROIs from DDSM dataset

<table>
<thead>
<tr>
<th>Segmentation Method</th>
<th>Performance Measures</th>
<th>FCM</th>
<th>IFCM</th>
<th>sIFCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (%)</td>
<td></td>
<td>79</td>
<td>81</td>
<td>86</td>
</tr>
<tr>
<td>Sensitivity</td>
<td></td>
<td>0.78</td>
<td>0.8</td>
<td>0.84</td>
</tr>
<tr>
<td>Specificity</td>
<td></td>
<td>0.8</td>
<td>0.82</td>
<td>0.86</td>
</tr>
</tbody>
</table>

From Tables 3.7 and 3.8, it is observed that the proposed sIFCM segments the TP ROIs exactly than that of the FCM and IFCM clustering techniques. Moreover, the classification results are better for ROIs segmented using sIFCM due to the proper segmentation of the ROI, when compared to FCM and IFCM.

The classification performance of the proposed sIFCM is compared with the performance of FCM and IFCM for both MIAS and DDSM datasets and shown in the following figures.
The proposed sIFCM is compared with FCM and IFCM. Figures 3.10 and 3.11 show that, the proposed sIFCM performs better in terms of classification accuracy, sensitivity and specificity than that of FCM and IFCM. In FCM, the membership function is always not accurately defined due to personal error whereas in the IFCM and the proposed the sIFCM, the
personal error is removed by using hesitation margin. IFCM utilizes the distance between the data points and cluster centers in the spatial domain to compute the membership function. The pixels on an image are highly correlated, and this spatial information is an important characteristic that can be used to aid labeling the pixels. However, the spatial relationship between pixels is seldom utilized in IFCM. The proposed sIFCM incorporates the spatial information into the membership function to improve the segmentation results, owing to which, the new sIFCM technique is less sensitive to noise than other methods.

3.5 SUMMARY

This chapter has dealt with the development of a novel and automatic sIFCM segmentation method in mammogram. This sIFCM clustering method achieves a significantly high performance than the two state-of-the-art segmentation methods, namely FCM and IFCM clustering. This significant improvement in segmentation performance results in better classification of benign and malignant ROIs.