

International Journal of Computational Intelligence and Applications
© World Scientific Publishing Company

Rough Morphology Hybrid Approach for Mammography Image Classification and Prediction

Aboul ella Hassanien

*Information Technology Department, FCI, Cairo University
5 Ahamed Zewal Street, Orman, Giza, Egypt **

Ajith Abraham

*Center for Quantifiable Quality of Service in Communication Systems
Norwegian University of Science and Technology
O.S. Bragstads plass 2E, NO-7491 Trondheim, Norway†*

Received (received date)

Revised (revised date)

The objective of this research is to illustrate how rough sets can be successfully integrated with mathematical morphology and provide a more effective hybrid approach to resolve medical imaging problems. Hybridization of rough sets and mathematical morphology techniques has been applied to depict their ability to improve the classification of breast cancer images into two outcomes: malignant or benign cancer. Algorithms based on mathematical morphology are first applied to enhance the contrast of the whole original image; to extract the region of interest and to enhance the edges surrounding that region. Then, features are extracted characterizing the underlying texture of the regions of interest by using the gray-level co-occurrence matrix. The rough set approach to attribute reduction and rule generation is further presented. Finally, rough morphology is designed for discrimination of different regions of interest to test whether they represent malignant cancer or benign cancer. To evaluate performance of the presented rough morphology approach, we tested different mammogram images. The experimental results illustrate that the overall performance in locating optimal orientation offered by the proposed approach is high compared with other hybrid systems such as rough-neural and rough-fuzzy systems.

Keywords: Keyword1; keyword2; keyword3.

1. Introduction

Breast cancer is a deadly disease that adversely affects the lives of far too many people, primarily women. According to the National Cancer Institute¹³, each year about 180,000 women in the United States develop breast cancer, and about 48,000 lose their lives to this disease. It is also reported that a woman's lifetime risk of developing breast cancer is 1 in 8. The breast cancer screening tests performed on a regular basis play a crucial role in reducing the rate of mortality, especially among

*Email: a.hassanien@fci-cu.edu.eg,abo@cba.edu.kw

†Email: ajith.abraham@ieee.org, abraham.ajith@acm.org

2 *Authors' Names*

women ages 50 and older. The screening tests include digital mammography, clinical breast examination, breast self-examination, or a combination of the above. Digital mammography refers to the application of digital system techniques on digital Mammograms. Currently, digital mammography is one of the most promising cancer control strategies since the cause of breast cancer is still unknown^{12,36}.

Rough set theory, initially introduced by Zdzisław Pawlak in the early 1980s⁵⁶ and further developed over the last 20 years (see, *e.g.*,^{51,52,53,54,55}) provides an approach to approximation of sets, which leads to useful forms of granular computing. The basic idea underlying the rough set approach to information granulation is to discover to what extent a given set of objects (*e.g.*, pixel windows in an image) approximates another set of objects of interest. Objects are compared by considering their descriptions. Rough set theory offers a novel approach to manage uncertainty that has been used for the discovery of data dependencies, importance of features, patterns in sample data, feature space dimensionality reduction, and the classification of objects. Rough sets have proven useful for representation of vague regions in spatial data. One important result of rough set theory is the extractions of *if-then* rules from the information tables. Such rules have the potential to reveal new patterns in sample data. In addition, rough set methods can be used to classify new sample data based on what is already known. Unlike other computational intelligence techniques, rough set analysis requires no external parameters and uses only the information presented in given sample data. Rough set theory can be used to assess the completeness of sample data: if the data is incomplete, it suggests more information about the objects needed to be collected to build a good classification model. On the other hand, if the data is complete, rough sets can determine the minimum number of features needed for classification. In effect, the rough set approach leads to feature space dimensionality reduction. This is very important for applications where domain knowledge is very limited or data collection is very expensive and laborious because it makes sure that the data collected is just good enough to build a good classification model without sacrificing the accuracy of the classification model or wasting time and effort to gather extra information about the objects^{49,50,27,51}.

In recent years, rough set have been combined with various other intelligent techniques to provide more adaptive and more effective algorithms, for example, Hassanien and Selzak⁶ introduced a rough neural approach for rule generation and image classification. Algorithms based on fuzzy image processing were first applied to enhance the contrast of the whole original image; to extract the region of interest and to enhance the edges surrounding that region. Then, they extract features characterizing the underlying texture of the regions of interest by using the grey-level co-occurrence matrix. Then, the rough set approach to feature reduction and rule generation was presented. Finally, rough neural network is designed for discrimination of different regions of interest to test whether they represent malignant cancer

or benign cancer. Rough neural network was built from rough neurons, each of which can be viewed as a pair of sub-neurons, corresponding to the lower and upper bounds. They run tests over different mammogram images to evaluate performance of their rough neural approach.

Also defined in the literature are rough-fuzzy sets⁴⁴, which can be seen to be a particular case of fuzzy-rough sets. A rough-fuzzy set is a generalisation of a rough set derived from the approximation of a fuzzy set in a crisp approximation space. This corresponds to the case where only the decision attribute values are fuzzy; the conditional values are crisp. The lower and upper approximations indicate the extent to which objects belong to a target set. Mao et al.³² proposed a new fuzzy Hopfield-model net based on rough-set reasoning for the classification of multispectral images. The main purpose is to embed a rough-set learning scheme into the fuzzy Hopfield network to construct a classification system called a rough-fuzzy Hopfield net (RFHN). The classification system is a paradigm for the implementation of fuzzy logic and rough systems in neural network architecture. Instead of all the information in the image being fed into the neural network, the upper- and lower-bound grey levels, captured from a training vector in a multispectral image, are fed into a rough-fuzzy neuron in the RFHN. Therefore, only $2/N$ pixels are selected as the training samples if an N -dimensional multispectral image was used. In the simulation results, the proposed network not only reduces the consuming time but also reserves the classification performance see also [1].

Mathematical morphology was invented in the early 1960s by Georges Matheron and Jean Serra who worked on the automatic analysis of images occurring in mineralogy and petrography^{16,20,21}. Meanwhile the method has found immense applications in several other fields, including medical diagnostics, histology, industrial inspection, computer vision, and character recognition^{8,22}. Mathematical morphology examines the geometrical structure of an image by probing it with small patterns, called structuring elements, of varying size and shape, just the way a blind man explores the world with his fingers or a stick. This procedure results in nonlinear image operators which are well-suited to explore geometrical and topological structures. A succession of such operators is applied to an image in order to make certain features apparent, distinguishing meaningful information from irrelevant distortions, by reducing it to a sort of skeletonization.

In this paper, a rough-morphology approach for rule generation, image classification and prediction is presented. Hybridization of rough sets and mathematical morphology techniques has been applied to illustrate their ability to improve accuracy to classify breast cancer images into two outcomes: malignant cancer or benign cancer. Algorithms based on mathematical morphology are first applied to enhance the contrast of the whole original image; to extract the region of interest and to enhance the edges surrounding that region. Features characterizing the underlying

texture of the regions of interest are extracted by using the gray-level co-occurrence matrix. Rough morphology is designed for discrimination of different regions of interest to test whether they represent malignant cancer or benign cancer. To evaluate performance of the presented rough morphology approach, we run tests over different mammogram images.

Rest of the paper is organized as follows. Section 2 gives a brief introduction to rough sets and mathematical morphology used in our approach. Section 3 discusses the proposed Rough morphology hybrid approach in Mammography images in details. Experimental results are discussed in Section 4. The paper is concluded in Section 5.

2. Basic Concepts

2.1. *Rough Set theory*

Rough set theory^{49,27,51} is a fairly new intelligent technique for managing uncertainty that has been applied to the medical domain and is used for the discovery of data dependencies, evaluates the importance of attributes, discovers the patterns of data, reduces all redundant objects and attributes, seeks the minimum subset of attributes, recognize and classify objects in image processing. Moreover, it is being used for the extraction of rules from databases. Rough sets have proven useful for representation of vague regions in spatial data. One advantage of the rough set is the creation of readable *if-then* rules. Such rules have a potential to reveal new patterns in the data; furthermore, it also collectively functions as a classifier for unseen data sets. Unlike other computational intelligence techniques, rough set analysis requires no external parameters and uses only the information presented in the given data. One of the nice features of rough set theory is that it can tell whether the data is complete or not based on the data itself. If the data is incomplete, it suggests more information about the objects are required to be collected in order to build a good classification model. On the other hand, if the data is complete, rough set approach can determine whether there are more than enough or redundant information in the data and find the minimum data needed for classification model. This property of rough set approach is very important for applications where domain knowledge is very limited or data collection is very expensive/laborious because it makes sure the data collected is just good enough to build a good classification model without sacrificing the accuracy of the classification model or wasting time and effort to gather extra information about the objects^{49,27,51}.

In rough set theory, the data is collected in a table, called decision table. Rows of the decision table correspond to objects, and columns correspond to attributes. In the data set, we assume that the a set of examples with a class label to indicate the class to which each example belongs are given. We call the class label the decision attributes, the rest of the attributes the condition attributes. Rough set theory defines three regions based on the equivalent classes induced by the attribute values

lower approximation, upper approximation and boundary. Lower approximation contains all the objects, which are classified surely based on the data collected, and upper approximation contains all the objects which can be classified probably, while the boundary is the difference between the upper approximation and the lower approximation. So, we can define a rough set as any set defined through its lower and upper approximations. On the other hand, indiscernibility notion is fundamental to rough set theory. Informally, two objects in a decision table are indiscernible if one cannot distinguish between them on the basis of a given set of attributes. Hence, indiscernibility is a function of the set of attributes under consideration. For each set of attributes we can thus define a binary indiscernibility relation, which is a collection of pairs of objects that are indiscernible to each other. An indiscernibility relation partitions the set of cases or objects into a number of equivalence classes. An equivalence class of a particular object is simply the collection of objects that are indiscernible to the object in question. Here we provide an explanation of the basic framework of rough set theory, along with some of the key definitions. Reader's may consult ^{49,27,51} for more fundamental details on rough set theory and applications.

Mathematical Morphology

Mathematical morphology is a set algebra used to process and analyze data based on geometric shapes. The theory of mathematical morphology was introduced by Matheron ¹⁶ in 1974 and refined by Serra ²⁰ in the 1980's. The basic morphological operations are erosion and dilation. For binary signals, erosion is a Minkowski set subtraction (i.e., an intersection of set translations), and dilation is a Minkowski set addition (i.e., a union of set translations). These operators were extended to operate on non-binary signals ^{39,37,14,10}. There are two main types of morphological filter: Binary morphological filters and Grayscale morphology filters. Set processing filters accept binary input signals and give binary output signals, while function processing filters accept binary or non-binary functions as input and yield non-binary functions as output. The interpretation of binary signals as sets and non-binary signals as functions is straightforward.

2.1.1. Binary morphological filters

Let X denotes an m -dimensional set and B denote a compact k -dimensional set ($k < m$), and let y denote a point in k and z a point in m . The set X is a binary signal or image to be filtered, and the set B is called the structuring element of the morphological filter. Define the symmetric set $\check{B} = -z : z \in B$, which is a reflection of B about the origin. The translation of a set to a point z is denoted by a subscript; for example, the set B translated to the point z is B_z . The set processing morphological erosion and dilation are defined by:

$$Erosion : X \ominus \check{B} = (z : B_z \subseteq X) = \bigcap_{y \in B} X_y \quad (1)$$

$$\text{Dilation} : X \oplus \check{B} = \{z : B_z \cap X \neq \phi\} = \cup_{y \in B} X_y \quad (2)$$

The symbols \ominus and \oplus denote Minkowski subtraction and addition, respectively. Erosion and dilation are complementary operations, one shrinking the size of objects in an image and the other expanding them. However, erosion and dilation are not inverses of each other. Some objects are completely removed by erosion, and therefore cannot be restored by dilation. Likewise, dilation often joins nearby objects which erosion cannot then separate. The compound morphological operations formed by performing the complementary operators in sequence are the morphological operations opening and closing. Opening is defined as erosion followed by dilation, while closing is dilation followed by erosion. The structuring element used for the second operation is the reflection about the origin of the structuring element used for the first operation.

$$\text{Opening} : O(X, \check{B}) = ((X \ominus \check{B}) \oplus B) \quad (3)$$

$$\text{Closing} : C(X, \check{B}) = ((X \oplus \check{B}) \ominus B) \quad (4)$$

Examples of erosion, dilation, opening and closing of some simple discrete sets are depicted in Figure 1. As evident, erosion shrinks a set while dilation expands a set. In most applications of mathematical morphology, the structuring elements are symmetric about the origin, so that $B = \check{B}$. When this is the case, there is no distinction between erosion and Minkowski subtraction nor between dilation and Minkowski addition. The structuring element rotation between the erosion and dilation creates an "effective" structuring element for opening and closing that is symmetric about the origin, even if the original structuring element is asymmetric⁴³.

2.1.2. *Grayscale morphology filters*

Since most signal and image processing applications do not deal with binary data, mathematical morphology must be extended to non-binary signals (functions) to be widely useful. The extension performs by representing a function as an ordered set of binary signals^{38,39}. Therefore, in this Section, we will extend the definitions of erosion and dilation for grayscale operations.

Grayscale erosion darkens small bright areas, and very small bright areas like noise spikes or small spurs might be totally removed¹⁷. Working with a 3x3-structuring element as above, grayscale erosion is implemented as follows:

1. The numerical values of the point and its eight neighbors are evaluated
2. Minimal value of these nine values is found

(Rough Morphology Hybrid Approach for Mammography Image Classification and Prediction) 7

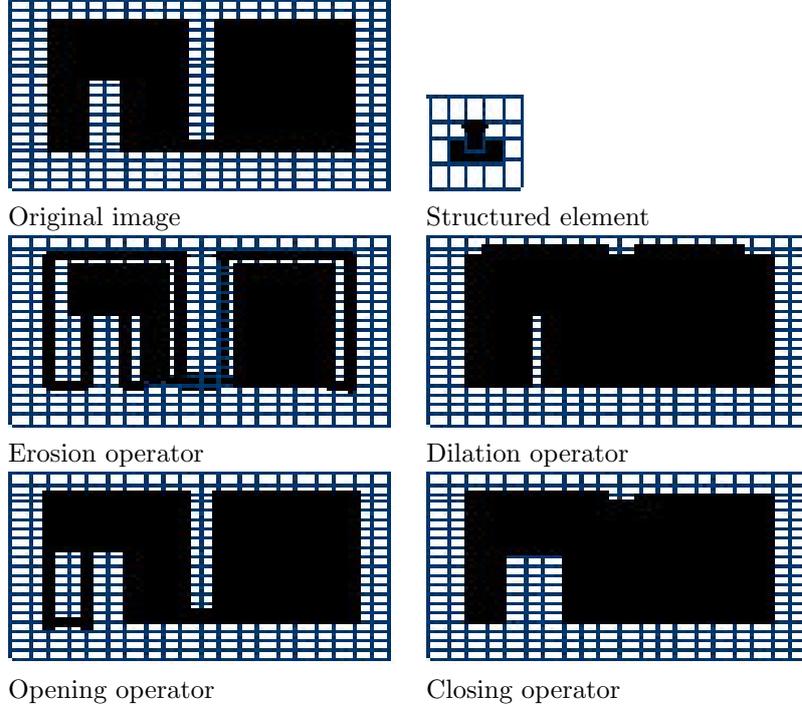


Figure 1: Binary morphological operators results

3. The new value of the point is set to the minimal value such that the erosion is defined as follows:

$$E_G(A, B) = \text{Min}_{[j,k] \in B} \{a[m + j, n + k] - b[j, k]\} \quad (5)$$

Grayscale dilation brightens small dark areas, and very small dark "holes" might be totally removed¹⁷. Working with a 3x3-structuring element as above, greyscale dilation is implemented as follows:

1. The numerical values of the point and its eight neighbors are evaluated.
2. Maximum value of these 9 values is found
3. The new value of the point is set to the maximal value.

The dilation is defined as follows:

$$D_G(A, B) = \text{Min}_{[j,k] \in B} \{a[m - j, n - k] - b[j, k]\} \quad (6)$$

First eroding the image with the structuring element, and then dilating the resulting image by the structuring element can achieve Greyscale opening¹⁷. The

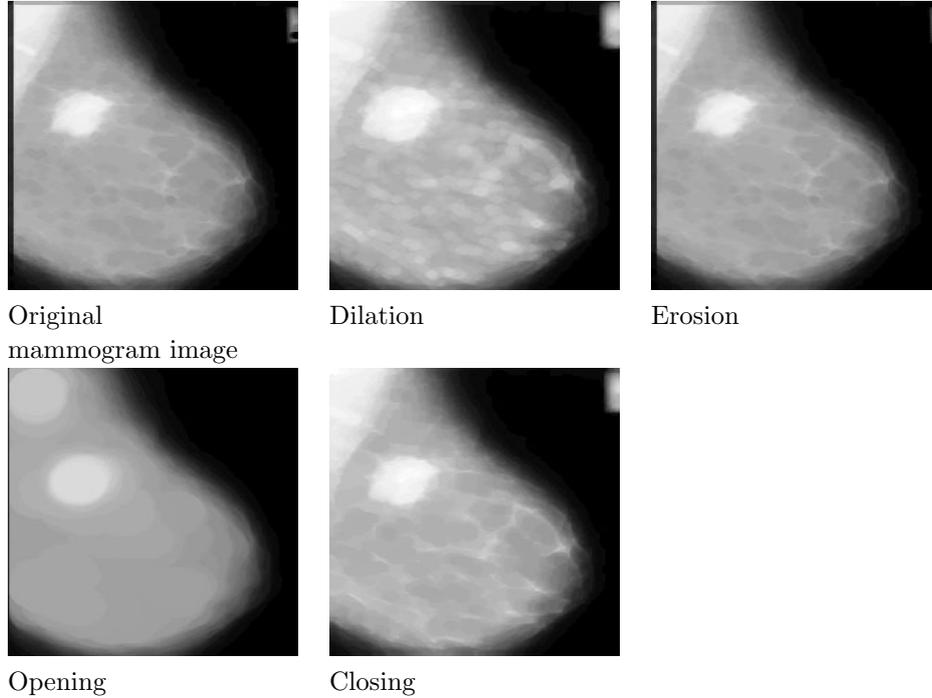


Figure 2: Illustration of Grayscale morphological operators

process darkens small bright areas, and may entirely remove very small bright spots like noise pikes. The opening process is defined as follows:

$$G_G(A, B) = D_G(E_G(A, B), B) \quad (7)$$

Grayscale closing is achieved by first dilating the image with the structuring element, and then eroding the resulting image by the structuring element. The process brightens small dark areas, and may entirely remove very small dark holes¹⁷. The closing process is defined as follows:

$$C_G(A, B) = E_G(D_G(A, B), B) \quad (8)$$

Examples of erosion, dilation, opening and closing operators of some real mammogram images are depicted in Figure 2. As evident, dilation brightens small dark areas, and very small dark *holes* might be totally removed. While erosion darkens small bright areas and very small bright areas like noise spikes or small spurs might be totally removed. Opening operator darkens small bright areas, and may entirely remove very small bright spots like noise pikes, while closing operator brightens small dark areas, and may entirely remove very small dark holes.

3. Rough morphology hybrid approach

Usually, the breast cancer detection system starts with the preprocessing stage, which includes digitization of the mammograms with different sampling and quantization rates. Then, the regions of interests selected from the digitized mammogram are enhanced. The segmentation process is designed to find suspicious areas and to separate the suspicious areas from the background that will be used for extracting features of suspicious regions. In the feature extraction and selection process, the features of suspicious areas will be extracted and selected, and suspicious regions will be classified into two classes: cancer or non cancer^{2,7,1,4}.

In this paper a rough morphology approach is adopted for designing a hybrid system for rule generation and image classification. The architecture of the proposed hybrid approach is illustrated in Figure 3. It is comprised of four fundamental building phases: pre-processing, feature extraction, rough set analysis and classification and prediction. In the first phase of the investigation, a pre-processing algorithms based on mathematical morphology are presented. It is adopted to improve the quality of the images and to make the feature extraction phase more reliable. Pre-processing stage involves enhancing mammogram images before a reasonable segmentation can be achieved as well as suppression noise in the segmented region of interest. Then, segmentation process involves grouping adjacent pixels with similar properties together to form connected regions. In the feature extraction phase, features extracted from the image regions are used to assign them to one of two classes: normal or abnormal. These stages have been added to provide a framework for the automatic analysis of the mammogram images, which is then evaluated using a pre-diagnosed image database. Rough set is used to classify the new image. The third phase is the rough set data analysis. It is done by computing the minimal number of necessary attributes, together with their significance, and generating the sets of rules. The last phase is the prediction of new objects, it is dependent on the type of generated rules and strategies based on the rough data analysis model. These four phases are described in detail in the following Section along with the steps involved and the characteristics feature for each phase.

3.1. *Pre-processing phase*

Mammography is a branch of radiology which could benefit greatly from the assimilation of digital imaging technologies. Diagnosing cancer tissues using digital mammograms is a time consuming task even for highly skilled radiologists because mammograms are low contrast, noisy images. Computerized enhancement techniques could be used to ensure optimum presentation of all digital clinical images as well as before a reasonable segmentation can be achieved. Image enhancement in medical computing is the use of computers to make an image clearer. This may be to aid the interpretation by humans or computers. Types of image enhancement include, noise reduction, edge enhancement and contrast enhancement. Contrast en-

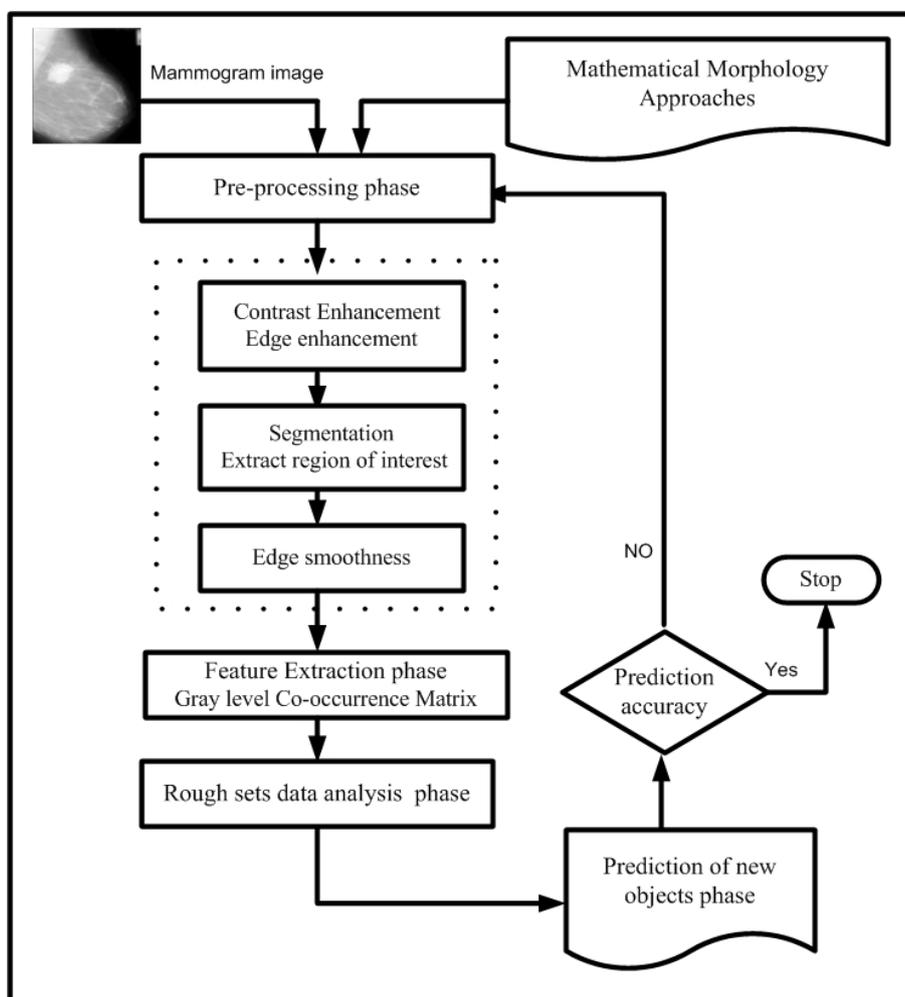


Figure 3: Rough Morphology Hybrid Approach

hancement is useful when an area of the image that is of particular importance has only subtle changes in pixel intensity. In these cases it may be difficult for the human eye to make out the structures clearly, especially if the image is being displayed on a low quality screen, or limited contrast like mammograms image. Applying the contrast enhancement filter will improve the readability of areas with subtle changes in contrast but will also destroy areas of the image where the intensity of the pixels is outside the range of intensities being enhanced. The proposed techniques based on mathematical morphology theory will treat this problem. Also enhancing the edges of the segmented regions of interest is necessary to improve the matching process. In this section, we introduce pre-processing techniques based on mathematical

morphology theory to enhance the contrast of the original mammogram image to help the segmentation process and enhancing the edges of the segmented regions of interest.

3.1.1. *Morphological contrast enhancement process*

The aim of contrast enhancement is to extract homogeneous areas of an image. The opening filter is used to enhance the original image by removing the bright component that does not fit within structuring element, while closing filter enhance the original image by removing the dark component that do not fit within structuring element. In this paper, we implement the open-close/close-open filter to enhance the contrast of the original images ²¹.

3.1.2. *Morphological edge enhancement process*

The human eye easily interprets edges and therefore edge enhancement can help a specialist to interpret an image. Edge enhancement techniques usually destroy other information in the image, such as color or intensity of the separate areas in the image. If this information is required it is possible to superimpose the edge-enhanced image over the top of the original image to give both the enhanced edges and the color or intensity information. Presence of noise, scattering and low-contrast makes the task even more difficult. One solution for this problem is mathematical morphology. Mathematical morphology is a powerful, shaped-based method often used in image processing.

Many image transformations can be derived from the basic morphological operators. Top-hat algorithm is a nice enhancement tool for edge enhancement that can be obtained from openings or closings operators. It has the property of enhancing peaks or valleys by applying respectively the opening or closing operator. The top-hat transform is defined as the difference between the original image and its opening. The opening of an image is the collection of foreground parts of an image that fit a particular structuring element. The key mechanism of the Top-hat transform under the opening operator is the local comparison of a shape, the structural element, with the object that will be transformed. If, when positioned at a given point, the structural element is included in the object then the whole structural element will appear in the result of the transformation, otherwise none of its points will appear.

After smoothing the image with the opening and closing operations, respectively, subtracting the original image yields the expected enhanced result. Due to the simplification properties of the closing operator, the top-hat has the property of enhancing valleys, that is, darker regions in the image, serving as a good background remover. The choice of the structuring element size is very important. If it is too small, the background removal is affected by the presence of random noise. If it is too large, it cannot reproduce the small-scale variations of the faint objects. Therefore, a good compromise has to be found.

3.1.3. Segmentation: Extracting the region of interest

The aim of the segmentation process is to extract the abnormal regions or regions with high probability of abnormality that is called Regions Of Interest (ROI) from the digital mammography image. The segmentation of lesion in mammographic images is known to be a difficult step toward automatic breast cancer diagnosis, due to the very low contrast of these images. Many different approaches to the problem of lesion segmentation can be found in the literature ^{48,35}. We applied the histogram thresholding algorithm, which uses intensity values to split the image domain into segmented objects regions of interest and background areas. We aim to detect abnormal and normal lesions in a mammogram by identifying suspicious regions of tissue, where a suspicious region is an abnormal region or a region with a high probability of abnormality. Here the segmentation process is applied on the result from the pre-processing step, morphological enhanced image, not on the original image. The main steps of the threshold segmentation algorithm are provided below.

Algorithm 1 Threshold segmentation algorithm

1 Input: original mammogram image

Output: Segmented region of interest

- 1: Apply morphological enhancement on original image;
 - 2: Get the histogram of this enhanced image
 - 3: Get threshold value T from the histogram
 - 4: Apply threshold segmentation as the following
 - 5: each pixel in the image
 - 6: **if** value of the pixel in the input image $>$ the threshold value T **then**
 - 7: assign this value to the corresponding pixel in output image
 - 8: Else
 - 9: set the value of corresponding pixel in output image with Zero
 - 10: **end if**
 - 11: The segmented image
-

3.1.4. Top-hat algorithm: edge surrounding enhancement

Morphological Top-hat transformation was proposed by Meyer ²⁵ is a nice enhancement tool for edge enhancement that can be obtained from openings or closings operators. It has the property of enhancing peaks or valleys by applying respectively the opening or closing operator. This approach allows the classification of regions of the image by size and height. The top-hat transformation is divided in white and black top-hat. In the case of dealing with clear regions the white top-hat is employed, and it is obtained as the subtraction between the original image

and the opening transformation. For dark regions, the black top-hat is used, this transformation is determined by the subtraction between the closing transformation and the original image²². The key mechanism of the Top-hat transform under the opening operator is the local comparison of a shape, the structural element, with the object that will be transformed. Figure 4 demonstrates the effect of the Top-hat transform using opening operator with a small 3x3 square-structuring element.



Figure 4: Top-hat transform using opening operator with a small 3x3 square structuring element

3.2. Feature extraction phase

Gray level co-occurrence matrix (GLCM) is one of the most known texture analysis methods that estimates image properties related to second-order statistics. Each entry (i, j) in GLCM corresponds to the number of occurrences of the pair of gray levels i and j which are a distance d apart in original image. In order to estimate the similarity between different gray level co-occurrence matrices, Haralick⁴⁶ proposed 14 statistical features extracted from them. To reduce the computational complexity, only some of these features were selected. In this paper, we adopt four features; energy, entropy, contrast and inverse difference moment. **Energy**, also called Angular Second Moment, is a measure of textural uniformity of an image. Energy reaches its highest value when gray level distribution has either a constant or a periodic form. A homogenous image contains very few dominant gray tone transitions, and therefore the normalized entry of the co-occurrence matrix for this image will have fewer entries of larger magnitude resulting in large value for energy feature. In contrast, if the normalized entry of the co-occurrence matrix contains a large number of small entries, the energy feature will have smaller value. The second feature is **entropy** which measures the disorder of an image and it achieves its largest value when all elements in normalized entry of the co-occurrence matrix are equal. When the image is not texturally uniform many GLCM elements have very small values, which implies that entropy is very large. Therefore, entropy is inversely proportional to GLCM energy. The third feature is **contrast**, which is a difference

moment of the normalized entry of the co-occurrence matrix and it measures the amount of local variations in an image. The last feature is the **inverse difference moment**. It measures image homogeneity. This parameter achieves its largest value when most of the occurrences in GLCM are concentrated near the main diagonal. Inverse difference moment is inversely proportional to GLCM contrast. For further reading, the reader may consult ⁵.

3.3. *Rough sets data analysis phase*

One way to construct a simple model computed from data, easier to understand and with more predictive power, is to create a set of minimal number of rules. Some condition values may be unnecessary in a decision rule produced directly from the database. Such values can then be eliminated to create a more comprehensible minimal rule preserving essential information. The rough set data analysis is illustrated in Figure 5, consists of four processes: (a) Discretizing the image features (i.e attributes) and create the decision table; (b) Significant of the attributes: Calculate the weights of the attributes; (c) Generate the reducts with minimal number of attributes; (d) Rule generation and rule accuracy: generate a list of rules, compute the overall accuracy of the generated rules. These four processes are described in detail in the following section along with the steps involved and the characteristics feature for each process.

3.3.1. *Rough Discretization process*

When dealing with attributes in concept image classification, it is obvious that they may have varying importance in the problem being considered. Their importance can be pre-assumed using auxiliary knowledge about the problem and expressed by properly chosen weights. However, in the case of using the rough set approach to concept classification, it avoids any additional information aside from what is included in the information table itself. Basically, the rough set approach tries to determine from the data available in the information table whether all the attributes are of the same strength and, if not, how they differ in respect of the classifier power. Therefore, some strategies for discretization of real valued features must be used when we need to apply learning strategies for data classification (*e.g.*, equal width and equal frequency intervals). It has been shown that the quality of learning algorithm is dependent on this strategy, which has been used for real-valued data discretization ¹⁵.

A pixel grey value in an image is always a numeric value. Many classification algorithms such as rough set theory, require that training data contain only discretized feature values. Otherwise, too many equivalent classes will be produced and the algorithms will be over sensitive to noise. To use such an algorithm when there are numeric-valued features, all numeric values must first be converted into discrete values-a process called discretization. This process is performed by dividing

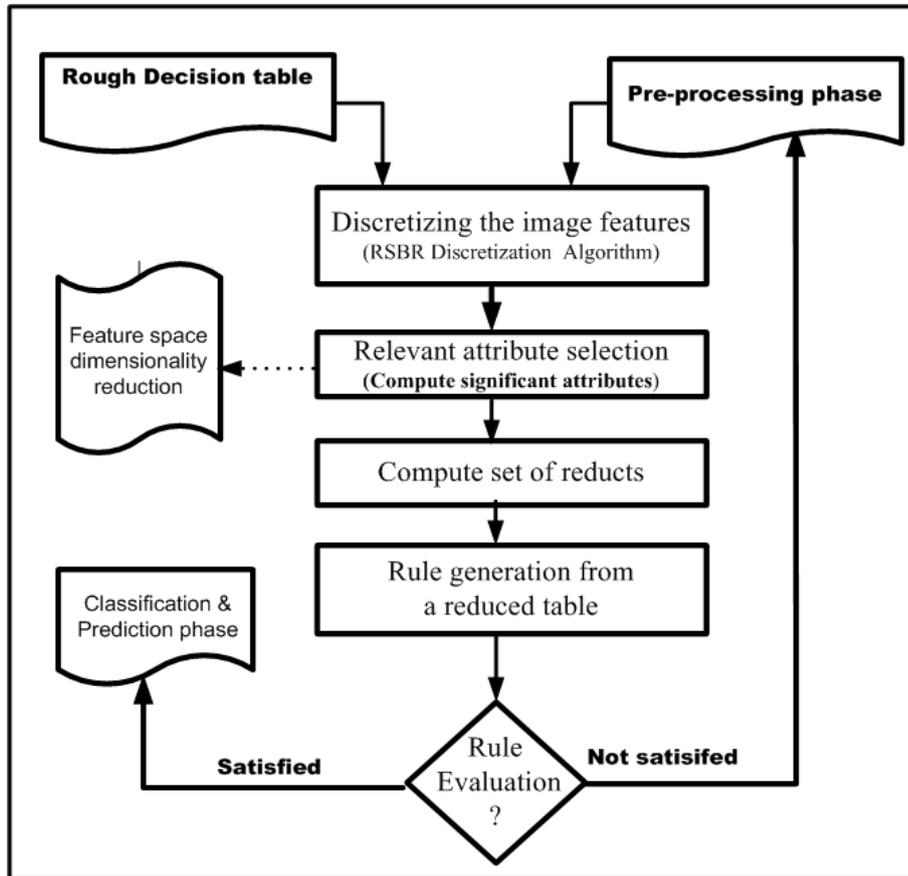


Figure 5: Rough Sets Data Analysis Phase

the values of a continuous attributes into a small number of intervals, where each interval is mapped to a discrete categorical, nominal, symbolic symbol. Discretization can significantly influence the effectiveness of a classification algorithm. There are two kinds of discretization methods: non-supervised and supervised. Since the former perform poorly in many situations because they ignore the class of the training examples, a typical supervised method, χ^2 -test also named ChiMerge algorithm is used to discretize pixel grey values. Another example is to use the rough set with boolean reasoning (RSBR) algorithm for the discretization of continuous-valued attributes proposed by Nguyen in 1997²⁶ (see, also,^{3,18,47}). The main advantage of RSBR is that it combines discretization of real valued attributes and classification. The main steps of the RSBR discretization algorithm is detailed below:

Algorithm 2 RSBR discretization Algorithm

Require: Information system table (S) with real valued attributes A_{ij} and n is the number of inter-values for each attribute.

Ensure: Information table (ST) with discretized real valued attribute.

- 1: **for** $A_{ij} \in S$ **do**
- 2: Define a set B of boolean variables in (9).

$$B = \left\{ \sum_{i=1}^n C_{ai}, \sum_{i=1}^n C_{bi}, \sum_{i=1}^n C_{ci}, \dots, \sum_{i=1}^n C_{ni} \right\}, \quad (9)$$

where $\sum_{i=1}^n C_{ai}$ correspond to a set of intervals defined on the variables of attributes a .

- 3: **end for**
- 4: Create a new information table S_{new} by using the set of intervals C_{ai}
- 5: Find the minimal subset of C_{ai} that discerns all the objects in the decision class D using (10).

$$\Upsilon^u = \wedge \{ \Phi(i, j) : d(x_i) \neq d(x_j) \}, \quad (10)$$

where $\Phi(i, j)$ is the number of minimal cuts that must be used to discern two different instances x_i and x_j in the information table.

Relevant attribute selection and reduction

The implemented feature extraction procedure relies on the texture, which is the main descriptor for all kinds of mammograms. Attribute reduction (feature selection) is a process of finding an optimal subset of all attributes according to some criterion so that the attribute subset are good enough to represent the classification relation of data. A good choice of attribute subset provided to a classifier can increase its accuracy, save the computational time, and simplify its results⁴². Attributes deleted in attribute reduction can be classified into two categories. One category contains irrelevant and redundant attributes that have no any classification ability. An irrelevant attribute does not affect classification in any way and a redundant feature does not add anything new to classification³⁴. The other category is noisy attributes. These attributes represent some classification ability, but this ability will disturb the prediction of true classification relation due to the effect of noise^{42,33}. Some researchers have proposed attribute reduction algorithms based on the rough set and a comprehensive overview can be found in^{45,30,31,9}

In general, rough set theory provides useful techniques to reduce irrelevant and redundant attributes from a large database with a lot of attributes^{45,30,31,9}. The dependency degree (or approximation quality, classification quality) and the information entropy⁶ are two most common attribute reduction measures in rough set theory. In this paper, we use the dependency degree measure to compute the significant features and measuring the effect of removing of the feature from the feature

(*Rough Morphology Hybrid Approach for Mammography Image Classification and Prediction*) 17

sets³.

3.3.2. *Computation of the reducts*

A reduced table can be seen as a rule set where each rule corresponds to one object of the table. The rule set can be generalized further by applying rough set value reduction method. Its main idea is to drop those redundant condition values of rules and unite those rules in the same class. Unlike most value reduction methods, which neglect the difference among the classification capabilities of condition attributes, we first remove values of those attributes that have less discrimination factors. Thus more redundant values can be reduced from decision table and more concise rules can be generated.

The computation of the reducts from a decision table is a way of selecting relevant features^{23,24}. It is a global method in the sense that the resultant reducts represent the minimal sets of features which are necessary to maintain the same classification accuracy given by the original and complete set of attributes. A straight manner for selecting relevant features is to assign a measure of relevance to each attribute and choose the attributes with higher values. Based on the reduct system, we generate the list of rules that will be used for building the classifier model for the new objects. In decision tables, there often exist conditional attributes that do not provide (almost) any additional information about the objects. So, we should remove those attributes since it reduces complexity and cost of decision process^{23,27,24}. A decision table may have more than one reduct. Anyone of them can be used to replace the original table. Finding all the reducts from a decision table is NP-complete. Fortunately, in applications, it is usually not necessary to find all of them – one or a few of them are sufficient enough. A natural question is, which reducts are the best. The selection depends on the optimality criterion associated with the attributes. If it is possible to assign a cost function to attributes, then the selection can be naturally based on the combined minimum cost criteria. In the absence of an attribute cost function, the only source of information to select the reduct is the contents of the table. In this paper, we adopt the criteria that the best reducts are the ones with the minimal number of attributes and – if there are more such reducts – with the least number of combinations of values of its attributes cf.⁶. In this paper we applied the reduct algorithm based on the degree of dependencies and the discrimination factors³. The main steps of the reduct generation algorithm are provided below:

Rule generation from a reduced table

The generated reducts are used to generate decision rules. The decision rule, at its left side, is a combination of values of attributes such that the set of (almost) all

Algorithm 3 Reduct Generation algorithm

Input: information table (ST) with discretized real valued attribute.

Output: reduct sets $R_{final} = \{r_1 \cup r_2 \cup \dots \cup r_n\}$

- 1: **for** each condition attributes $c \in C$ **do**
 - 2: Compute the correlation factor between c and the decisions attributes D
 - 3: **if** the correlation factor > 0 **then**
 - 4: Set c as relevant attributes.
 - 5: **end if**
 - 6: **end for**
 - 7: Divide the set of relevant attribute into a different variable sets.
 - 8: **for** each variable sets **do**
 - 9: Compute the dependency degree and compute the classification quality
 - 10: Let the set with high classification accuracy and high dependency as an initial reduct set.
 - 11: **end for**
 - 12: **for** each attribute in the reduct set **do**
 - 13: Calculate the degree of dependencies between the decisions attribute and that attribute.
 - 14: Merge the attributes produced in previous step with the rest of conditional attributes
 - 15: Calculate the discrimination factors for each combination to find the highest discrimination factors
 - 16: Add the highest discrimination factors combination to the final reduct set.
 - 17: **end for**
 - 18: **repeat**
 - statements 12
 - 19: **until** all attributes in initial reduct set is processed
-

objects matching this combination have the decision value given at the rule's rough side. The rule derived from reducts can be used to classify the data. The set of rules is referred to as a classifier and can be used to classify new and unseen data. The main steps of the rule generation and classification algorithm are provided below:

The quality of rules is related to the corresponding reduct(s). We are especially interested in generating rules, which cover largest parts of the universe U . Covering U with more general rules implies smaller size of a rule set.

3.4. Classification and prediction phase

Classification and prediction is the last phase of our proposed approach. Figure 6 illustrates the classification scheme for a construction of particular image classification and prediction algorithm. To transform a reduct into a rule, one only has to bind the condition feature values of the object class from which the reduct orig-

Algorithm 4 Rule generationInput: reduct sets $R_{final} = \{r_1 \cup r_2 \cup \dots \cup r_n\}$

Output: Set of rules

-
- 1: **for** each reduct r **do**
 - 2: **for** each correspondence object x **do**
 - 3: Contract the decision rule $(c_1 = v_1 \wedge c_2 = v_2 \wedge \dots \wedge c_n = v_n) \longrightarrow d = u$
 - 4: Scan the reduct r over an object x
 - 5: Construct $(c_i, 1 \leq i \leq n)$
 - 6: **for** every $c \in C$ **do**
 - 7: Assign the value v to the correspondence attribute a
 - 8: **end for**
 - 9: Construct a decision attribute d
 - 10: Assign the value u to the correspondence decision attribute d
 - 11: **end for**
 - 12: **end for**
-

inated to the corresponding features of the reduct. Then, to complete the rule, a decision part comprising the resulting part of the rule is added. This is done in the same way as for the condition features. To classify objects, which has never been seen before, rules generated from a training set will be used. These rules represent the actual classifier. This classifier is used to predict to which classes new objects are attached. The nearest matching rule is determined as the one whose condition part differs from the feature vector of re-image by the minimum number of features. When there is more than one matching rule, we use a voting mechanism to choose the decision value. Every matched rule contributes votes to its decision value, which are equal to the t times number of objects matched by the rule. The votes are added and the decision with the largest number of votes is chosen as the correct class. Quality measures associated with decision rules can be used to eliminate some of the decision rules (cf. ¹).

The global strength defined in ²³ for rule negotiation is a rational number in $[0, 1]$ representing the importance of the sets of decision rules relative to the considered tested object. Let us assume that $T = (U, A \cup \{d\})$ is a given decision table, u_t is a test object, $Rul(X_j)$ is the set of all calculated basic decision rules for T , classifying objects to the decision class $X_j(v_d^j = v_d)$, $MRul(X_j, u_t) \subseteq Rul(X_j)$ is the set of all decision rules from $Rul(X_j)$ matching tested object u_t . The global strength of decision rule set $MRul(X_j, u_t)$ is defined by the following form ²³:

$$MRul(X_j, u_t) = \frac{\left| \bigcup_{r \in MRul(X_j, u_t)} |Pred(r)|_A \cap |d = v_d^j|_A \right|}{\left| |d = v_d^j|_A \right|}.$$

Measure of strengths of rules defined above is applied in constructing classification algorithm. To classify a new case, rules are first selected matching the new

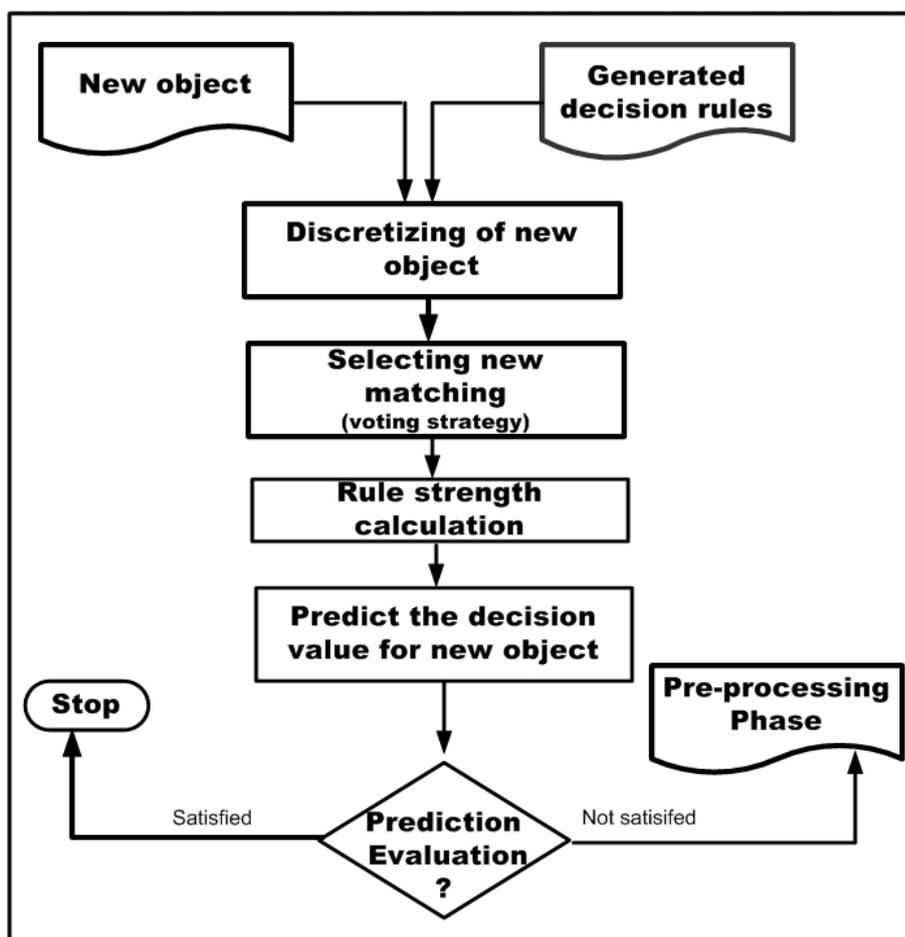


Figure 6: Rough Classification and Prediction Scheme

object. The strength of the selected rule sets is calculated for any decision class, and then the decision class with maximal strength is selected, with the new object being classified to this class.

4. Experiment Results and Discussion

To evaluate the visual performance of the algorithm some images from the Mammographic Image Analysis Society (MIAS) database were selected ²⁹.

In our experiments, all the test data were abnormal breast mammogram images and the breast abnormality was spiculated lesion. Figure 7 depicts the enhancement results of the original mammogram image. Where Figure 7.a represents the original

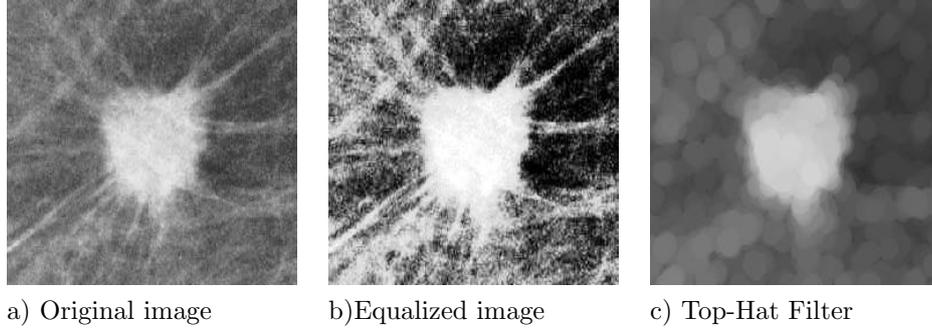


Figure 7: Pre-processing enhancement results

image, Figure 7.b shows the equalized histogram enhancement results and Figure 7.c illustrates the morphological enhancement results. From Figure 7, it is evident that the contrast enhancement of the original mammogram image by morphological filter exhibited the best results compared with histogram equalization results. The morphological filter algorithm is preferred in the pre-processing stage since it has two main advantages:

- 1) Increase the contrast between target area and background area.
- 2) Decrease the overlap between target and the gray level distribution.

Figure 8 represents the results of the segmentation algorithm. Figure 8.a represents the original image, Figure 8.b shows the enhanced image using morphological enhancement, Figure 8.c depicts the segmented region of interest, using threshold segmentation approach, Figure 8.d illustrate the result of the region of interest after applying closing mathematical morphology operator to enhance the segmented regions of interest by filling the small dark areas within the segmented image, where these small holes make some drawbacks in matching process and Figure 8.e shows the result of the region of interest after applying opening mathematical morphology operator to enhance the segmented regions of interest by removing the small bright areas within the segmented image and enhance the edges of ROI, which help in matching process.

4.0.1. *Generated rule evaluation*

To evaluate the sets of induced rules, the rule importance measure is taken into account. It is defined as follows:

$$Importance_{rule} = \frac{\tau_r}{\rho_r}, \quad (11)$$

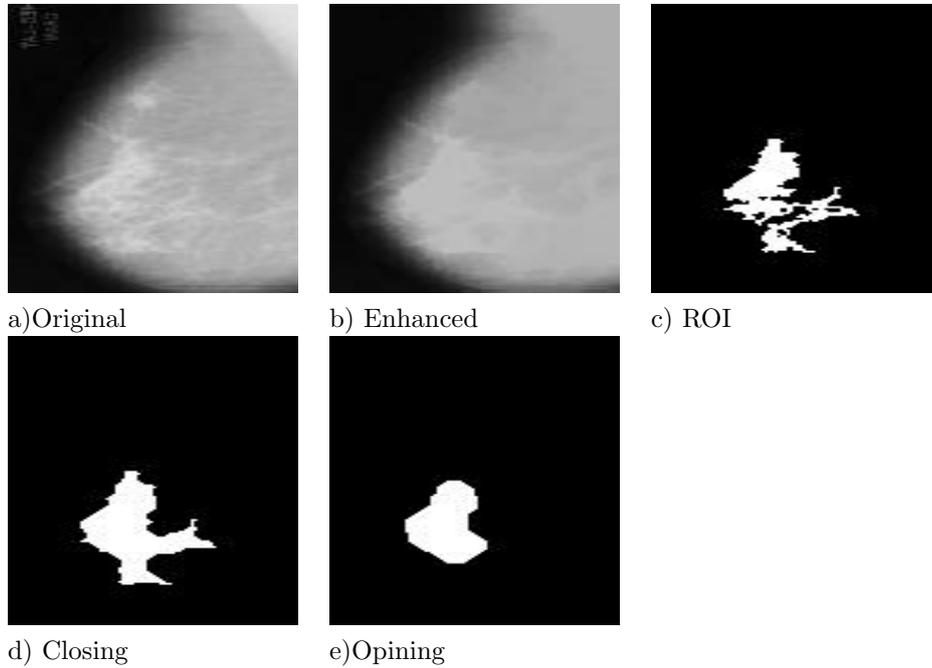


Figure 8: Segmented results: Region of interest

where τ_r is the number of times a rule appears in all reducts and ρ_r is the number of reduct sets.

Figure 9 depicts the number of generated rules before and after pruning process. We can observe that the number of generated rules for all algorithms is very large. Actually, it is greater than the number of objects. It makes classification unacceptably slow. Therefore, it is necessary to prune the rules during their generation. On the other hand, rule-based classification process associated with neural networks is not easy to explain as rules that are meaningful to the user. Moreover, in the decision tree, rough- neural, and rough- fuzzy required more robust features to improve the performance of the classifier and in particular for the generated number of rules.

4.0.2. *Classification accuracy evaluation*

Classification experiments were performed using 153 samples of the breast cancer mammogram information. It contains the computational statistical data for building the rough classifier. In our implementation, we divide the 153 samples of data into two sets of samples such that a first set is used for testing the model that has been developed from the training set. To assess the performance of the model on the test data set, in this paper we used the sensitivity and specificity measures to estimate the classifier accuracy for rough morphology approach. The classifier accuracy is

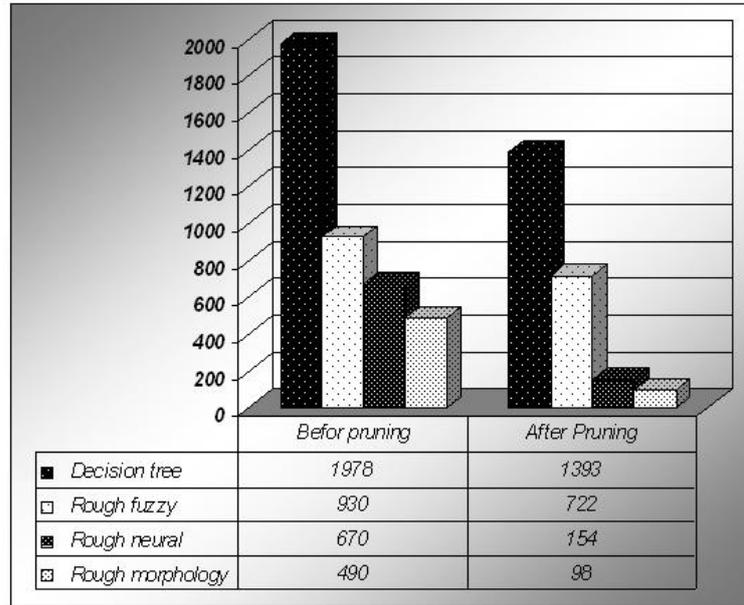


Figure 9. Generated number of rules before and after pruning

computed based on the following form:

$$Accuracy = \frac{T_p}{(pos + neg)} + \frac{T_n}{(pos + neg)} \quad (12)$$

Where T_p is the number of true positive (samples that are correctly classified as such), pos is the number of positive samples, T_n is the number of true negatives (samples that are correctly classified as such) and neg is the number of negative samples.

Figure 10 illustrates the overall classification accuracy in terms of sensitivity and specificity compared with Decision tree ⁷, Discriminant analysis ³, Rough-neural ⁶ and Rough-fuzzy ¹. Empirical results reveal that the proposed rough-morphology approach performs better than the other classifiers.

5. Conclusions and future works

A fully integrated and hybrid combination of various computational intelligence technologies in pattern recognition and, in particular, medical imaging problems has become one of the most promising avenues in image processing research. From the perspective of rough set, further explorations into possible hybridization of rough set with other computational and non-computational technologies are necessary to build a more complete picture of rough-based applications in medical imaging.

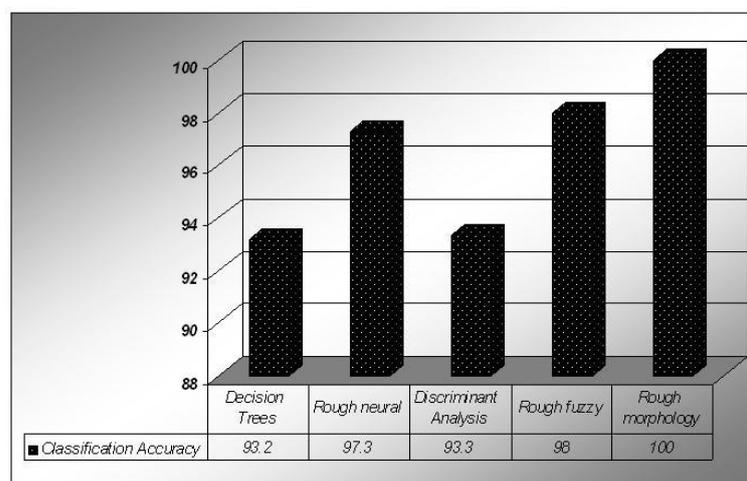


Fig. 10 Classification accuracy of rough morphology compared with Decision tree-, Discriminant Analysis-, Rough neural- and Rough fuzzy- classifiers.

What can be said at this point is that the rough set approach paves the way for new and interesting avenues of research in medical imaging and represents an important challenge for computational intelligence researchers.

This paper presented a hybrid rough-morphology approach for image classification and prediction of digital mammogram images. Experiment results illustrate that the overall performance in locating optimal orientation offered by the proposed approach is high compared with other hybrid systems such as rough-neural and rough-fuzzy systems.

Our future works will focus on building a fully integrated rough-morphology system in pattern recognition and, in particular, medical imaging problems. Bloch⁴¹ illustrated that there is a tight relationship between rough set and mathematical morphology, that gives evidence of the increasing interest of the image analysis researchers towards these theories. Based on the observation that rough set and mathematical morphology are both using dual operators sharing similar properties, Bloch⁴¹ also investigated more closely the links existing between both the domains. Equivalence between some morphological operators and rough sets has been established from either a relation, or a pair of dual operators or a neighborhood system. Author suggested some extensions using morphological thinning and thickening using algebraic operators.

Recently, many successful work towered this issue has been addressed and discussed. For example, Yang et al.⁴⁰ proposed a method called membership-grade operator for binary image processing. Based on the same structural element, the op-

(*Rough Morphology Hybrid Approach for Mammography Image Classification and Prediction*) 25

erator achieves multiform dilations, erosions and edge extractions by proper choice of membership grade obtained by binary image. Applied to noisy image, the operator can achieve efficient processing and noise reduction, and good effect is obtained by image simulation. Authors also analyzed the characteristics of the membership-grade operator by comparing it with the morphology operator.

Skowron and Polkowski¹¹ proposed a method called analytical morphology for data filtering. The method is based on some ideas of rough set theory and mathematical morphology. Mathematical morphology makes an essential use of geometric structure of objects while the aim of their method is to provide tools for data filtering when there is no directly available geometric structure in the data set.

Complexity of morphological operations makes it desirable to propose a theoretical scheme of an approximate morphological calculus within a general paradigm of soft computing. Polkowski²⁸ proposed a scheme based on ideas from rough set theory. In the proposed scheme, the underlying space of pixels is partitioned into disjoint classes by means of some primitive features and morphological operations are performed on classes, which allows for compression of image data¹⁹.

References

1. A.E. Hassanien, Fuzzy-rough hybrid scheme for breast cancer detection, in *Image and Computer Vision Journal*, **25(2)** February, (2007) 172-183.
2. A.E. Hassanien, Ali J.M., and N. Hajime, Detection of spiculated masses in Mammograms based on fuzzy image processing, in *7th International Conference on Artificial Intelligence and Soft Computing, ICAISC2004*, Zakopane, Poland, June 7-11. LNAI, Springer, **3070** 1002-1007.
3. A.E. Hassanien, M.E. Abdelhafez, H.S. Own, Rough Sets Data Analysis in Knowledge Discovery: A Case of Kuwaiti Diabetic Children Patients, *Advances in Fuzzy System*, (2008), Article ID 528461, 13 pages, doi:10.1155/2008/528461.
4. A.E. Hassanien, Classification and feature selection of breast cancer data based on decision tree algorithm, in *International Journal of Studies in Informatics and Control*, **12(1)**(2003) 33-39.
5. A.E. Hassanien and J.M. Ali, Rough Set Approach for Classification of Breast Cancer Mammogram Images, in *Lecture Notes in Computer Science* **2955** (2006) 224-231.
6. A.E. Hassanien and D. Selzak, Rough Neural Intelligent Approach for Image Classification: A Case of Patients with Suspected Breast Cancer, in *International Journal of Hybrid Intelligent Systems* **3(4)** (2006) 205-218.
7. A.E. Hassanien, Classification and feature selection of breast cancer data based on decision tree algorithm, in *International Journal of Studies in Informatics and Control*, **12(1)** (2003) 33-39.
8. A.E. Hassanien and E.H. Tarek Abed, Digital Mammography Image Analysis System Based on Mathematical Morphology, in *IEEE computer society 7th International Conference On Intelligent Engineering Systems INES 2003*, March 4-6 Assiut-Luxor, Egypt, 141-147.
9. A. Verikas and M. Bacauskiene, Feature selection with neural networks., in *Pattern Recognition Letters* **23** (2002)13231335.
10. A.J. Candas, U. M. Neto, and E. C. Filho, A Mathematical Morphology Approach to the Star/Galaxy Characterization Problem, in *Journal of the Brazilian Computer*

Society 3(3)(1997)

11. A. Skowron and L. Polkowski, Analytical Morphology: Mathematical Morphology of Decision Tables, in *Fundamenta Informaticae*, **27(2/3)** (1996) 255-271.
12. B.K. Verma, and J. Zakos A Computer-Aided Diagnosis System For Digital Mammograms Based On Fuzzy-Neural And Feature Extraction Techniques, in *IEEE Transactions on Information Technology in Biomedicine*, **5** (1999) 46-54.
13. Cancer Net Home Page of the National Cancer Institute <http://biomed.nus.sg>, 2002.
14. E. R. Dougherty *Mathematical Morphology in Image Processing*, Marcel Dekker, New York (1993).
15. J. Dougherty, R. Kohavi, and M. Sahami, Supervised and Unsupervised Discretization of Continuous Features, in *Proc. of the XII International conference on Machine Learning*, (1995) 294-301.
16. G. Matheron, *Random Sets and Integral Geometry* (New York Wiley, (1974)).
17. L. Miller and N. Ramsey, The detection of abnormal masses by non-linear multiscale analysis, in *3rd International Workshop on Digital Mammography* (1996) 335-340.
18. H.S. Nguyen, On efficient handling of continuous attributes in large data bases , in *Fundamenta Informaticae*, **48(1)** 1, (2001) 6181.
19. J.G. Stell., Relations in Mathematical Morphology with Applications to Graphs and Rough Sets, in *Lecture Notes in Computer Science, Springer Berlin / Heidelberg*, **4736/2007 Spatial Information Theory**, (2007) 438-454.
20. J. Serra, *Image Analysis and Mathematical Morphology 1* (London, Academic press, (1982).
21. J. Serra, *Image Analysis and Mathematical Morphology, 2* heoretical Advances, London, Academic (1988).
22. J. D. Mendiola-Santibaez, I. R. Terol-Villalobos, G. Herrera-Ruiz, and A. Fernandez-Bouzas, Morphological contrast measure and contrast enhancement: One application to the segmentation of brain MRI, in *Signal Processing Journal*, **87(9)** September (2007) 2125-2150.
23. J. Bazan, H. S. Nguyen, S. H. Nguyen, P. Synak, and J. Wróblewski, Rough Set Algorithms in Classification Problem, in *Polkowski, L., Tsumoto, S., Lin, T. Y. (eds.), Rough Set Methods and Applications. Physica Verlag* (2000) 49-88.
24. J. A. Starzyk, N. Dale, and K. Sturtz, A mathematical foundation for improved reduct generation in information systems, in *Knowledge and Information Systems Journal*, Springer (2000) 131-147.
25. F. Meyer, *Cytologie quantitative et morphologie mathmatique*, These de Docteur-Ingénieur, Ecole des Mines de Paris, 1979.
26. H.S. Nguyen, *iscretization of Real-Value Attributes, Boolean Reasoning Approach*, Ph.D. Thesis, supervisor: A. Skowron, Warsaw University, Poland, 1997.
27. L. Polkowski, *Rough Sets: Mathematical Foundations*, (Physica-Verlag (2003)).
28. L. Polkowski, Approximate mathematical morphology: rough set approach, <http://citeseer.ist.psu.edu/326941.html> (pdf file)
29. <http://www.wiau.man.ac.uk/services/MIAS/MIASmini.htm> (2003) The Mammographic Image Analysis Society: Mini Mammography Database.
30. N. Zhong, J. Dong and S. Ohsuga, Using rough sets with Heuristics for feature selection, in *Journal of Intelligent Information Systems* **16** (2001) 199214.
31. M. Beynon, Reducts within the variable precision rough sets model: A further investigation, in *European Journal of Operational Research* **134** (2001) 592-605.
32. C.W. Mao, S.H. Liu and J.-S. Lin, Classification of multispectral images through a rough-fuzzy neural network, *Optical Engineering*, **43(1)** (2004) 103-112.
33. M. Last, A. Kandel and O. Maimon, Information-theoretic algorithm for feature se-

(*Rough Morphology Hybrid Approach for Mammography Image Classification and Prediction*) 27

- lection, in *Pattern Recognition Letters* **22** (2001) 799811.
34. M. Dash and H. Liu, Feature selection for classification, in *Intelligent Data Analysis* **1** (1997) 131156.
 35. N. Pal and S. Pal, A review on image segmentation techniques, in *Pattern Recognition* **1(26)** 1277-1294.
 36. S. L. Kok, Computer Interpretation of X-ray Mammography, *Department of Engineering Science, University of Oxford* <http://www.ios.wcl2.ee.upatras.gr> (1993).
 37. P. Maragos and R. W. Schafer, Morphological filters-Part I: Their set-theoretic analysis and relations to linear shift-invariant filters, in *IEEE Trans. Acoust. Speech, Signal Process* **35(8)** 1153-1169.
 38. S. R. Sternberg, Grayscale morphology, in *Comp. Vision, Graphics, Image Process.* **35** (1986) 333-355.
 39. P. Soille, *Morphological Image Analysis, Principles and Applications*, (Springer (1999)).
 40. G.Q. Yang, L.H. Jiang and Y. Li., Application of Rough Sets in Binary Morphology, in *International Conference on Machine Learning and Cybernetics* (2006) 3446-3449.
 41. I. Bloch, On links between mathematical morphology and rough sets, in *Pattern Recognition* **33(9)**, (2000) 1487-1496.
 42. R. Li, and Z. Wanga, Mining classification rules using rough sets and neural networks, in *European Journal of Operational Research* **157(2)** September (2004) 439-448.
 43. R. M. Haralick, S. R. Sternberg, and X. Zhurag, Image Analysis Using Mathematical Morphology, in *IEEE trans on PAMI* **9(4)** 532-550.
 44. P. Srinivasan, M.E. Ruiz, D.H. Kraft, and J. Chen, Vocabulary mining for information retrieval: rough sets and fuzzy sets, *Information Processing & Management*, **37(1)** (1998) 1538.
 45. R.W. Swiniarski and A. Skowron, Rough set methods in feature selection and recognition, in *Pattern Recognition Letters* **24** (2003) pp. 833849.
 46. R.M. Haralick, Statistical and structural approaches to texture, in *Proceeding of the IEEE* **67(5)**, May (1979).
 47. Q. Zhang, An Approach to Rough Set Decomposition of Incomplete Information Systems, in *2nd IEEE Conf. on Industrial Electronics and Applications*, (2007) 2455-2460.
 48. R. Haralick and L. Shapiro, Image segmentation techniques, *CVGIP*, **29** (1995) 100-132.
 49. Pawlak Z., *Rough sets – Theoretical aspects of reasoning about data*, Kluwer, 1991.
 50. Z. Pawlak, J. Grzymala-Busse, R. Slowinski, and W. Ziarko, Rough Sets, in *Communications of the ACM* **8(11)** (1995) 88-95.
 51. Z. Pawlak, Rough sets, in *International J. Comp. Inform. Science*, **11**, (1982) 341-356.
 52. Z. Pawlak, Rough classification, in *International Journal of Man-Machine Studies*, **20(5)**, (1984) 469-483.
 53. Z. Pawlak, and A. Skowron, Rudiments of rough sets, *Information Sciences*, **177(1)** (2-7) 3-27.
 54. Z. Pawlak and A. Skowron, Rough sets: Some extensions, *Information Sciences. An International Journal, Elsevier*, **177(1)** (2007) 28-40.
 55. Z. Pawlak and A. Skowron, Rough sets and Boolean reasoning, *Information Sciences. An International Journal, Elsevier*, **177(1)** (2007) 41-73.
 56. Z. Pawlak, *Classification of Objects by Means of Attributes*, Institute for Computer Science, Polish Academy of Sciences, (1981) **Report 429**.