

A NOVEL APPROACH FOR AUTOMATIC DETECTION AND CLASSIFICATION OF SUSPICIOUS LESIONS IN BREAST ULTRASOUND IMAGES

Behnam Karimi and Adam Krzyżak

*Department of Computer Science and Software Engineering, Concordia University
Montreal, QC H3G 1M8, Canada*

Abstract

In this research, a new method for automatic detection and classification of suspected breast cancer lesions using ultrasound images is proposed. In this fully automated method, de-noising using fuzzy logic and correlation among ultrasound images taken from different angles is used. Feature selection using combination of sequential backward search, sequential forward search and distance-based methods is obtained. A new segmentation method based on automatic selection of seed points and region growing is proposed and classification of lesions into two malignant and benign classes using combination of AdaBoost, Artificial Neural Network and Fuzzy Support Vector Machine classifiers and majority voting is implemented.

1 Introduction

Interpretation of breast ultrasound images is a critical step in diagnosing breast cancer. Upon analyzing ultrasound images the radiologist determines whether to send a patient for biopsy. There are many challenges in interpretation of ultrasound images. Sometimes even experienced radiologists have difficulties to decide if a lesion is suspicious for cancer. If a radiologist misses cancer (false negative), then it can have serious consequences because the patient will not seek medical treatment and cancer can spread.

Interpretation of breast ultrasound images is a challenging task because ultrasound images (specially breast ultrasound images) are very noisy. This makes the interpretation very difficult as sometimes normal breast tissues are considered as part of the lesion and vice versa. The other reason that radiologists have sometimes difficulties analyzing breast ultrasound images is shadowing. Shadowing is not part of the normal breast tissue or lesion but it is an artifact present in ultrasound images. The shadows are sometimes mistakenly considered as

part of lesions and make the analysis very difficult.

A lot of research has been done in recent years and many algorithms have been proposed to automate detection of breast cancer. An algorithm proposed in [30] uniquely combines histogram equalization in preprocessing stage with hybrid filtering, multifractal analysis, thresholding segmentation, and a rule-based approach in fully automated regions of interest (ROI) labeling as shown in Figure 1.

The proposed method is able to very accurately label most of the lesions. It performs best with malignant lesions where it is able to identify correctly (90%) of them and its worst performance being the task of identification of fibroadenomas (77.59%). It appears that even by using Hybrid Filtering and Multifractal Processing, the accuracy of fibroadenomas detection is not very high. It is suspected that noise and shadowing in the images lead to low accuracy of the method.

Another approach introduced in [10] uses a bilateral subtraction technique to reduce false positives in mass candidate regions detected by detec-

tion scheme for the whole breast ultrasound images. By using the method, the performance of the CAD system is improved. It was found that the bilateral subtraction technique could reduce false positives effectively. This technique is based on the fact that normal left and right breasts of the same subject exhibit architectural symmetry. This method is based on using symmetrical features in both breasts. This is a useful tool used by radiologists to interpret ultrasound images. Even if there is a region like mass, the region is classified as normal if a similar region appears in the other breast in same position. This method uses this feature to reduce false positives. The method involves (1) image feature extraction, (2) registration of bilateral breasts, and (3) reduction of false positives. It removes 67.3% of false positives, but needs more improvements. It looks like the accuracy of the system can be improved by employing a better pre-processing technique for noise and shadow removal.

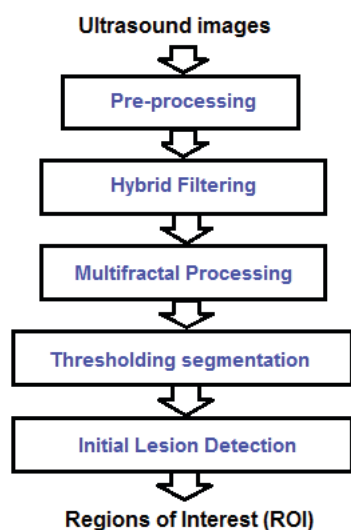


Figure 1. Fully automated ROI labelling system

Another method was proposed in [20] that uses speckle features of automated breast ultrasound images (ABUS). The ABUS images of 147 pathologically proven breast masses (76 benign and 71 malignant cases) were used. For each mass, a volume of interest (VOI) was cropped to define the tumor area and the average number of speckle pixels within a VOI was calculated. In addition, first-order and second-order statistical analysis of the speckle pixels was used to quantify the information of gray-level distributions and the spatial relations among the pixels. Receiver operating characteristic curve

analysis was used to evaluate the performance. It achieves the accuracy of 84.4%. The performance indices of the speckle features were comparable to the performance indices of the morphological features, which include shape and ellipse-fitting features. Although the accuracy is not ideal, it can be improved by combining speckle features with morphological and texture features.

The structure of the paper is as follows. The proposed system is summarized in Section 2. Detailed description of different components of the proposed system is provided in Section 3. It comprises compounding, segmentation, feature extraction and feature selection and classification. Simulation results on real data are presented in Section 4.

2 Proposed computer-aided diagnosis system

Figure 2 illustrates the conceptual diagram of the system proposed in this paper. It consists of the following modules.

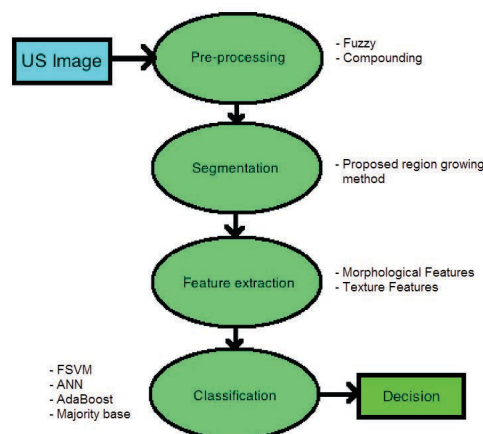


Figure 2. Proposed system

Pre-processing: Removes noise from ultrasound images and makes them ready for segmentation.

Segmentation: Determines the boundary of the suspected lesion(s).

Feature extraction and selection: Identifies the best features for classification purpose.

Classification: Classifies the lesion into different

classes to determine if the lesion is benign or malignant.

Pre-processing is the most important stage of the system. We need to remove noise and shadows as much as possible in order to better distinguish lesions from breast tissues. Because of the fuzzy nature of ultrasound images, we propose to use fuzzy logic for pre-processing and we performed several experiments validating our approach. We also realized that radiologists often look at a lesions from different angles. The reason is that sometimes due to shadowing and noise the lesion cannot be clearly recognized. Based on that finding, we suspected that correlation of ultrasound images from different angles could give us more information and eliminate some of the unwanted noise and shadows. We introduced a method of compounding ultrasound images from different angles and performed experiments to see if compounding can reduce noise and shadows while preserving important information in ultrasound images.

For feature selection, we used combination of some well-known feature selection algorithms such as Sequential Forward Search, Sequential Backward Search and Mutual Information (MI). We also extracted some texture features namely autocovariance, SGLDM, GLDM, BDIP, BVLC and NGTDM [2]. Selection of a subset of morphological features and extraction of texture features are described later in Section 4.

3 Pre-processing

Ultrasound images are usually significantly deteriorated by noise because of various sources of interferences and other phenomena. The noise usually appears as bright and dark spots and is called speckle noise. It obscures fine details, degrades quality of image and makes it difficult to detect low-contrast lesions. Speckle noise is often undesirable, as it makes it difficult to interpret the lesions and to choose the proper diagnosis. Thus in a computerized system for detection of ultrasound images, pre-processing to eliminate the noise is an important component of the system [10, 20, 8]. We introduce a novel approach for pre-processing of breast ultrasound images that is based on combination of fuzzy logic and compounding.

3.1 Compounding and correlation of images

Real time compounding of ultrasound images has been investigated for long time. Special equipment has been designed to perform compounding of ultrasound images from different angles. But despite of that conventional ultrasound is still being used [6]. We can still take advantage of compounding when using conventional ultrasound.

Although usage of ultrasound imaging is beneficial and it is noninvasive technique there are some restrictions in its application due to the physical nature of imaging. Some of the most important issues are

- High systematic noise in ultrasound images.
- Due to anisotropic resolution in ultrasound images and due to differences in the propagation velocity of sound waves in different types of tissues the geometric representation of objects is strongly dependent on the angle of insonification.
- Some artifacts like shadowing may hamper clear delineation of lesions.

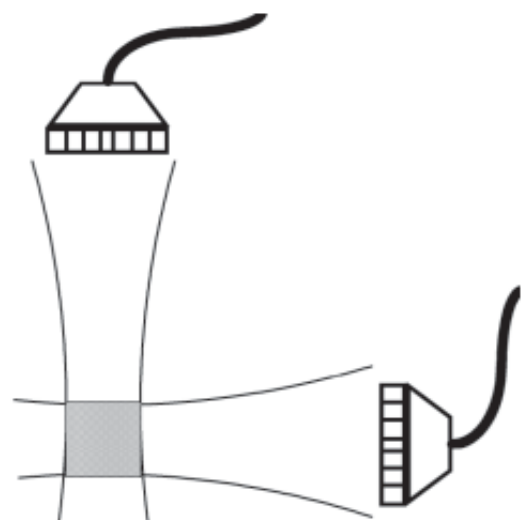


Figure 3. An ultrasound transducer is rotated fully around the female breast to acquire data from multiple angles (borrowed from [25])

Because of the issues mentioned above, ultrasound imaging is an interactive process that requires lots of experience to capture and correctly interpret ultrasound images. Findings in the ultrasound images are often not reproducible and vary among different interpreters.

The limitation mentioned above might be overcome by considering correlation between images from different angles (i.e. multiple viewing angles all around the breast). The concept is known as Full Angle Spatial Compounding (FASC) [26, 25]. When the ultrasound is being performed, the investigator usually moves the transducer around the female breasts to capture images from different angles. The correlation among those images could give more information and can make the interpretation easier. The method of capturing images from different angles is illustrated in Figure 3.

This method has capability of improving diagnosis for the following reasons:

- As noise is uncorrelated in images taken at different angles, it will be reduced by using this technique.
- Compound images exhibit an isotropic resolution which is a combination of the axial and lateral resolution of the individual images.
- Shadowing is suppressed because of varying angle of insonification.
- Structures which cause specular reflection are imaged and delineated in the compound image.

3.2 Fuzzy logic pre-processing

The main problem with most of the segmentation methods is that they are sensitive to noise and noise is unavoidable in ultrasound images. Another problem with some of the segmentation methods is that they produce different results for images from different types of sonographic machines. These problems suggest using a method capable of clearly separating background and foreground in ultrasound images. By doing so, the segmentation is expected to be more accurate. The uncertainty in identification lesions and lesions boundaries suggest use of fuzzy logic.

The detection of structures is crucial for the diagnosis of a vast number of illnesses including

breast cancer in breast ultrasound images. As ultrasound images are blurred by nature, have little contrast and are immersed in noise, most standard techniques of digital image processing do not yield good results for these images. Fuzzy logic that uses both global and local information has the ability to enhance fine details of the ultrasound images and is thus an approach of choice for low-contrast ultrasound images as their details cannot be obtained easily. The obtained enhancement allow to better distinguish the background from the actual image improving final detection of cancer.

An interval-valued fuzzy set states that the membership degree of every element to the set is given by a closed subinterval of interval $[0, 1]$. The type 2 fuzzy sets concept was introduced by Zadeh [31, 32] as a generalization of an ordinary fuzzy set. The membership degree of an element to a type 2 fuzzy set is a fuzzy set in $[0, 1]$.

An interval type 2 fuzzy set \bar{A} in U is defined by

$$\bar{A} = \{(u, A(u), \mu_u(x)) | u \in U, A(u) \in L([0, 1])\}, \quad (1)$$

where $A(u) = [\underline{A}(u), \bar{A}(u)]$ is a membership function; i.e., a closed subinterval of $[0, 1]$. Function $\mu(x)$ represents the fuzzy set associated with the element $u \in U$ obtained when x is within $[0, 1]$; $\mu_u(x)$ is defined as follows

$$F(x) = \begin{cases} a & \text{if } \underline{A}(u) \leq x \leq \bar{A}(u) \\ 0 & \text{otherwise.} \end{cases}$$

Sahba et al [22] proposed fuzzy rules for image enhancement, in which fuzzy rules such as the following ones have been used:

- IF pixel does not belong to the object, THEN leave it unchanged.
- IF pixel belongs to the breast object AND is dark, THEN make it darker.
- IF pixel belongs to the breast object AND is gray, THEN make it dark.
- IF pixel belongs to the breast object AND is bright, THEN make it brighter.

The degree of membership of each pixel to the object is a function of its distance to the central

point of the object or the inside of an initial/coarse segment. For initial segment, we use the region growing algorithm proposed in [28]. The initial segment is required to decide if a pixel belongs to the lesion. The main idea of enhancement is to eliminate noise in images and enhance gray levels of the selected area (regional contrast enhancement). In this method, each pixel is fuzzyfied depending on its intensity with a membership function that is constructed taking into account the mean level of gray of the surroundings and the position of the selected point. The fuzzy membership function that is used in this research is shown in Figure 4.

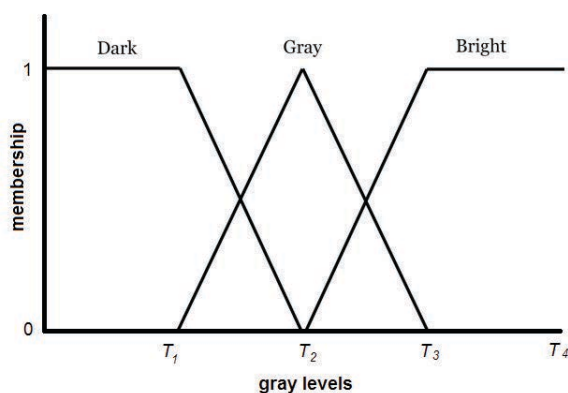


Figure 4. Membership functions for input gray level values

In Figure 4, T_4 is the brightest gray level in the image. For simplicity we used $T_1 = 25$, $T_2 = 50$ and $T_3 = 80$.

In this research, this method has been applied to breast ultrasound images to eliminate the unwanted noise. The idea is to map the image space to a fuzzy space using fuzzy rules and then apply segmentation techniques to detect lesions.

3.3 Feature extraction

The following geometrical features are extracted from ultrasound images.

- **Perimeter.** The Perimeter feature represents the length of the tumor perimeter. As malignant tumors usually have irregular shapes, a large tumor perimeter is associated with the likelihood that a tumor is malignant.
- **Area.** The Area feature is the area of a breast tumor. Malignant tumors frequently have a large area compared with benign tumors.

- **NSPD (number of substantial protuberances and depressions).** The NSPD feature can be utilized to calculate the level of boundary irregularity.
- **LI (lobulation index).** According to the definition for a concave point from the NSPD, the lobe region enclosed by a lesion contour and a line connected by any two adjacent concave points can be obtained. Usually, a malignant tumor has a larger LI than does a benign one.
- **ENC (elliptic-normalized circumference).** The angle of inclination for each tumor, with respect to the $x-y$ coordinate plane, can be obtained by using the second order moment.
- **ENS (elliptic-normalized skeleton).** The skeleton of a tumor region expresses a set S , and ENS is defined as the sum of the skeleton points in S . When a tumor has a twisted boundary, the skeleton is also complex. A malignant lesion always has a twisted boundary and generates a large ENS.
- **LS Ratio (long axis to short axis ratio).** The LS Ratio is the length ratio of the major (long) axis and minor (short) axis of the equivalent ellipse defined in the ENC feature.
- **Aspect Ratio.** The Aspect Ratio is the ratio of a tumor's depth and width. If a tumor depth exceeds its width, the Aspect Ratio is greater than 1 and the tumor has a high probability of being malignant.

$$\text{FormFactor} = \frac{4\pi \times \text{Area}}{\text{Perimeter}^2}.$$

When Form Factor is close to 1 then tumor is nearly round.

- **Roundness.**

$$\text{Roundness} = \frac{4 \times \text{Area}}{\pi \times \text{MaxDiameter}^2},$$

where Max Diameter denotes the length of the major axis from the equivalent ellipse of a tumor.

- **Solidity.**

$$\text{Solidity} = \frac{\text{Area}}{\text{ConvexArea}},$$

where Convex Area is the area of the convex hull of a tumor. When Solidity is close to 0, the tumor is malignant.

– **Convexity.**

$$Convexity = \frac{ConvexPerimeter}{Perimeter},$$

where Convex Perimeter is the perimeter of the convex hull of a tumor.

– **Extent.**

$$Extent = \frac{Area}{BoundingRectangle},$$

where Bounding Rectangle is the smallest rectangle containing the tumor.

– **TCA Ratio.** The TCA Ratio (tumor area to convex area ratio) is defined as:

$$TCARatio = \frac{Area}{ConvexArea}$$

– **TEP Ratio (tumor perimeter to ellipse perimeter ratio).** The TEP Ratio is the ratio of a tumor perimeter and the corresponding ellipse perimeter. The major and minor axes of the corresponding ellipse are calculated based on the proportion of width to depth of a tumor to acquire the same area for the ellipse and tumor.

– **TEP Difference (difference between tumor perimeter and ellipse perimeter).** The TEP Difference is defined as the difference between tumor perimeter and the corresponding ellipse perimeter.

– **TCP Ratio (tumor perimeter to circle perimeter ratio).** The TCP Ratio is the ratio of a tumor perimeter and the corresponding circle perimeter, the corresponding circle having the same area as the tumor.

– **TCP Difference (difference between tumor perimeter and circle perimeter).** The TCP Difference is defined as the difference between the tumor perimeter and the corresponding circle perimeter, the corresponding circle having the same area as the tumor.

– **AP Ratio (area to perimeter ratio).** The AP Ratio is the ratio of the area and the perimeter of a tumor.

– **Thickness of the wall.** If the wall of the mass is thick, there is a better chance that it is cancerous. If it is thin, it is more possible that it is a cyst rather than a malignant tumor.

Chen et al [2] studied several texture features in ultrasound images. These features included BDIP (Block difference of inverse probabilities), 2D normalized auto-covariance coefficients, SGLDM (Spatial gray-level dependence matrices), GLDM (Gray-level difference matrix) and NGTDM (Neighborhood gray-tone difference matrix). After extraction of features, PCA is applied to reduce the dimensionality. After applying PCA, the study considered all the possible combinations of texture features and ranks all possible combinations to extract the best features. Table 1 shows the result of the study that selects seven texture features for classification. The accuracy of this method is reported in a range of 65-84%.

Table 1. Texture features selected by Chen et al [2]
- A: 7x7 auto-covariance matrix; B: SGLDM; C: GLDM; D: BDIP; E: BVLC; F: NGTDM

Rank	Feature set
1	AD
2	ADE
3	AEF
4	ABCF
5	A
6	ADF
7	ACDE

3.4 Feature selection

In this research we used mutual information technique that ranks the features based on the discriminatory power. Then we used a selection algorithm to select a sub-set of those features. Due to limitations of the Sequential Forward Search (SFS) and the Sequential Backward Search (SBS) methods, a combination of those methods are used to make sure no important feature is eliminated by any of those methods. To guarantee that SFS and SBS converge to the same solution, we must ensure that features already selected by SFS are not removed by SBS and features already removed by SBS are not selected by SFS. In order to achieve that, every time SFS attempts to add a new features, we check if it has been removed by SBS. If it has been removed by SBS then we attempt to make it the second best feature.

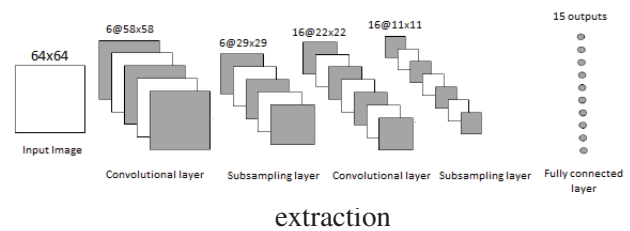
Our feature selection algorithm (MI, SFS and SBS) is used to select a subset of morphological

features shown in Table 2 and [2] selects a subset of texture features shown in Table 1. The order they appear is based on the rank of the features, i.e. Roundness is the best feature in morphological feature set and AD is the best feature in texture feature set.

Table 2. Selected morphological and texture features

Rank	Morphological Feature
1	Roundness
2	Solidity
3	Convexity
4	TCA Ratio
5	Perimeter
6	Area
7	NSPD
8	Aspect Ratio

We identified a subset of morphological features and texture features for classification of breast lesions. Selection of these features depends on a classifier and classification stage depends on these selected features. One method to extract features independent of a classifier is a Convolutional Neural Network (CNN) [15, 14]. [13] uses CNN to automatically extract features from handwritten digits. This is a feed-forward neural network that extracts topological features of the image by doing weighted convolutions and subsampling. The method is trained like a normal neural network using back propagation. This method uses several convolutional and subsampling layers which alternate (i.e. one convolutional layer is followed by the subsampling layer and vice-versa). Convolutional layer extracts elementary features from the image. It is organized in planes, also called feature maps, of simple units called neurons. It uses a 5×5 convolutional window that forms a unit in the input image or in the previous layer. A trainable weight is assigned to each connection, as it is done in normal neural networks. A convolutional layer consists of several feature maps. In order to apply CNN we converted the images in our database to the size of 64×64 . We used a convolutional window of 7×7 to perform the transformation [9]. Fig. 5 illustrates the architecture of the proposed system.



In subsampling layer, we have the same number of feature maps from the previous convolutional layer but half the number of rows and columns. Each unit j is connected to a 2×2 receptive field and we compute the average of its four inputs y_i , multiply it by a trainable weight w_j and add a trainable bias b_j to obtain the activity level v_j as follows:

$$v_j = w_j \frac{\sum_{i=1}^4 y_i}{4} + b_j.$$

3.5 Segmentation

A region growing method for segmentation of ultrasound images from [28] is adapted to our problem. This method applies the co-occurrence matrix features and gray level run-length features for identifying the seed point for given liver ultrasound images. We have implemented the method and ran it against our breast ultrasound images. We used fuzzy method to de-noise our images before applying the method. In region growing method the seed points are not selected correctly in some of the cases. An example is shown in Figure 6.

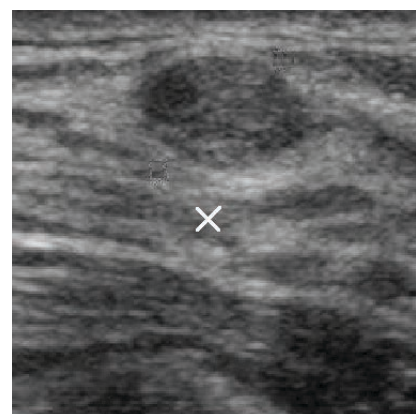


Figure 6. Incorrect seed point placement using method in [28]

The reason for this inaccurate selection of seed point is the fact that there are other noise areas in the

image that are considered as lesions. To overcome this problem we first identified the regions in the image, removed the noise region and rank the remaining regions based on importance and select the main region of interest. In order to achieve this, we first calculate all the local minimums of the image histogram. Then we need to find a good threshold to separate the lesion from the background. This threshold should be one of the local minimums. We used well-known Otsu thresholding method to achieve this [18]. This is a parameter free thresholding technique which maximizes the inter-class variance. It is interesting to observe that Otsu's method is more accurate in segmenting into two classes (foreground and background).

After finding the proper threshold we binarize and reverse ultrasound images using the threshold (lesions become white and background become black). Examples of binarized images are shown in Figure 7.

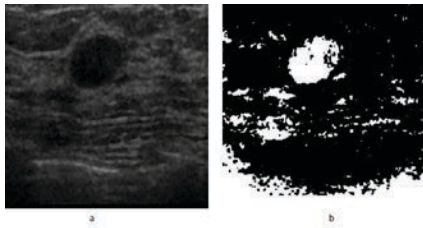


Figure 7. a) Original image b) Binarized image

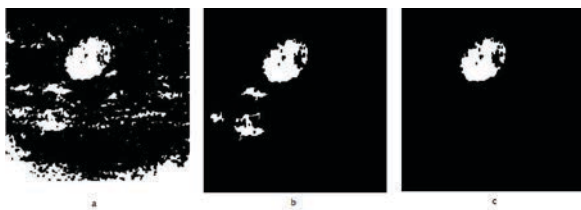


Figure 8. a) Original binarized image b) Binarized image after deleting boundary-connected region c) Winning region

In Figure 8 the left regions are either not connected with the boundary or they have an intersection with the image center window. We use the following score formula to rank each remaining region. The one with the highest score is considered

as the lesion region.

$$S_n = \frac{\sqrt{Area}}{dis(C_n, C_0).var(C_n)}, n = 1, \dots, k$$

where k is the number of regions, $Area$ is the number of pixels in the region, C_n is the center of the region, C_0 is the center of the image, and $var(C_n)$ is the variance of a small circular region centered at C_n .

Now that we identified the winning region, we can select the seed point. To select a seed point, we use a simple approach. Let us consider the minimum rectangle containing the winning region $[x_{min}, x_{max}; y_{min}, y_{max}]$. In most cases the center of the minimum rectangle could be considered as a seed point. However in some cases that the region shape is irregular, center point might be outside of the lesion. For those cases we can consider

$$x_{seed} = \frac{(x_{min} + x_{max})}{2},$$

$$y_{seed} = \{\forall y | (x_{seed}, y) \in lesion\ region\}.$$

Now that we have identified our seed point, we use the region growing method in [28] to complete the segmentation process.

The gold standard in segmentation is the segmentation proposed by a physician. This means that segmentation has to be performed and verified by a physician in order to validate the segmentation produced by the algorithm. As it is practically impossible to ask a physician to verify the outputs produced by several methods, we rely on existing research and our own experience to validate the results obtained by the automatic system.

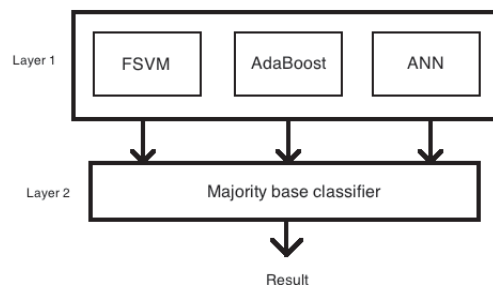
Recently the state-of-the-art segmentation method has been proposed by Shan, Cheng and Wang [23]. This method is based on neutrosophic I-means clustering and it outperforms other ultrasound image segmentation techniques to date, see [17], [33] and [16]. We introduce a new segmentation method combining region growing method and neutrosophic I-means clustering method. In Table 3 we compared region growing method, neutrosophic I-means clustering and our approach on our breast ultrasound images database. It is clear that our method performs best.

Table 3. Comparison of different segmentation methods

Method	Accuracy %
Region growing in [28]	92.50
l-means clustering [23]	93.75
Proposed method	97.50

3.6 Classification

We used several classification algorithms and found out that the best classifiers for classification of ultrasound images are Adaboost, Artificial Neural Network (ANN), Support Vector Machine (SVM) and Fuzzy Support Vector Machine (FSVM). FSVM outperforms other classifiers. Also AdaBoost and ANN are a very strong classifiers but we do not have any data on the comparison of AdaBoost with other classifiers. In our proposed method, we first let Adaboost, FSVM and ANN to produce their decisions and then we use a majority vote classifier to combine the outputs of the three classifiers. A conceptual model of our classifier is shown in Figure 10.

**Figure 10.** Proposed classifier - FSVM, AdaBoost, ANN classifiers and Majority vote classifier

For majority vote classifier, we used majority voting, see [12]. Assume that the *label* outputs of the classifiers are given as c -dimensional binary vectors $[d_{i,1}, \dots, d_{i,c}]^T \in \{0, 1\}^c, i = 1, \dots, L$, where $d_{i,j} = 1$ if D_i labels \mathbf{x} in ω_j , and 0 otherwise. The majority vote results in an ensemble decision for class ω_k if

$$\sum_{i=1}^L d_{i,k} = \max_{j \in \{1, \dots, c\}} \sum_{i=1}^L d_{i,j}. \quad (2)$$

Rule (2) is often called majority rule. It coin-

cides with the simple majority (50 percent of the votes+1) in case of two classes ($c = 2$).

4 Results

The methods used in our experiments are presented in Table 4. Please note that for segmentation we use the proposed region growing method. We also used 8 morphological features and 7 texture features shown in Table 2. The results of classification experiments are shown in Table 5.

Table 4. Experiments with different combinations of segmentation and classification algorithms. F: Fuzzy, C: Compounding

#	Pre-processing	Classifier
1	None	ANN
2	None	AdaBoost
3	None	FSVM
4	None	Proposed classifier
5	Fuzzy	ANN
6	Fuzzy	AdaBoost
7	Fuzzy	FSVM
8	F & C	Proposed classifier
9	F & C	ANN
10	F & C	AdaBoost
11	F & C	FSVM
12	F & C	Proposed classifier

Table 5. Comparison of our classifiers with state-of-the-art classifiers. Table 4 shows the methods used for each experiment. TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative. Highlighted rows are the result of our proposed classifier

#	TP	TN	FP	FN	Accuracy%
1	54	16	5	5	87.50
2	54	16	4	6	87.50
3	56	17	3	4	91.25
4	57	18	2	3	93.75
5	56	16	3	5	90.00
6	57	16	2	5	91.25
7	57	18	2	1	93.75
8	58	18	1	3	95.00
9	57	16	2	5	91.25
10	56	20	3	1	95.00
11	57	20	2	1	96.25
12	59	20	0	1	98.75

We also investigated the performance of the Convolutional Neural Network as features extractor. For segmentation, we used the proposed region growing method. The methods used in our experiments are described in Table 6 and the results of experiments are shown in Table 7. It appears that features extracted by the CNN do not improve the overall accuracy of the system.

Table 6. Experiments with different combinations of methods at each stage of the CAD system using CNN for feature extraction.



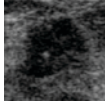
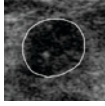
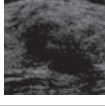
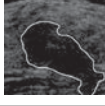
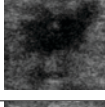

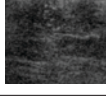

#	Pre-processing	Classifier
1	None	ANN
2	None	AdaBoost
3	None	FSVM
4	None	Proposed classifier
5	Fuzzy	ANN
6	Fuzzy	AdaBoost
7	Fuzzy	FSVM
8	F & C	
9	F & C	ANN
10	F & C	AdaBoost
11	F & C	FSVM
12	F & C	Proposed classifier

Table 7. Results of comparing our proposed classifiers with state-of-the-art classifiers using CNN for feature extraction. Table 6 shows the methods used for each experiment. TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative. Highlighted rows are the result of our proposed classifier

#	TP	TN	FP	FN	Accuracy%
1	52	14	7	7	82.50
2	52	14	7	7	82.50
3	54	14	5	7	85.00
4	56	18	3	3	92.50
5	53	15	6	6	85.00
6	55	18	4	3	91.25
7	57	17	2	4	92.50
8	57	19	2	2	95.00
9	54	16	5	5	87.50
10	55	18	4	3	91.25
11	57	18	2	3	91.25
12	57	19	2	2	95.00

Table 8 shows sample output of original ultrasound images, segmented images and the result of our proposed classifier.

Table 8. Sample output of ultrasound images, segmented images and the result of our classifier. CO: Classifier Output, GS: Gold Standard

Image	Segmented	CO	GS
		Malignant	Malignant
		Benign	Benign
		Malignant	Malignant
		Malignant	Malignant
		Malignant	Benign

5 Conclusions

The main problem with processing of ultrasound images is speckle noise and shadowing. These two issues present big challenge for designing automating system for detection of suspicious lesions in ultrasound images. To overcome some of the problems, we proposed a fully automated system for detection of breast ultrasound images. To remove the noise, we used fuzzy logic and compounding methods at pre-processing stage of our system. We were able to improve the quality of ultrasound images before segmentation, therefore we were able improve performance of the system. We have also identified a subset of features which worked well in classification. As experimenting with different combinations of features were almost practically impossible, we extracted most significant features for the purpose of classification. We implemented novel techniques for segmentation and classification. Our results show that our proposed system performs better than the state-of-the-art CAD systems for classification of ultrasound images.

In future we plan on experimenting with different sets of features and to apply our method to a database of color ultrasound images. We also intend to build an expert system for the diagnosis stage of our automated CAD system. The expert system will attempt to minimize the effort by a physician for diagnosis of suspicious lesions. We would also like to investigate the impact of texture features on the performance of the system.

References

- [1] R. Chang, W. Wu, W. Moon, and D. Chen. Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors. *Breast Cancer Research and Treatment*, 89(2):179–185, 2005.
- [2] D. Chen, Y. Huang, and S. Lin. Computer-aided diagnosis with textural features for breast lesions in sonograms. *Computerized Medical Imaging and Graphics*, 35:220–226, 2011.
- [3] Z. Dokur and T. Olmez. Segmentation of ultrasound images by using a hybrid neural network. *Pattern Recognition Letters*, 23(14):1825–1836, 2002.
- [4] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. Wiley, New York, 2001.
- [5] D. Eigen, J. Rolfe, R. Fergus, and Y. LeCun. Understanding deep architectures using a recursive convolutional network. *ArXiv e-prints*, 2013.
- [6] R. Entekin, P. Jackson, J. Jago, and B. Porter. Real time spatial compound imaging in breast ultrasound: technology and early clinical experience. *Medicamundi*, 43(3), 1990.
- [7] W. Gomez, W. Pereira, and A. Infantosi. Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound. *IEEE Transactions on Medical Imaging*, 31, 2012.
- [8] S. Gupta, R. Chauhan, and S. Sexena. Robust non-homomorphic approach for speckle reduction in medical ultrasound images. *Medical and Biological Engineering and Computing*, 43:189–195, 2005.
- [9] F. J. Huang and Y. LeCun. Large-scale learning with svm and convolutional nets for generic object categorization. *Proceedings of the 2006 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1:284–291, 2006.
- [10] K. Ikedo, Y. F. D., T. Hara, H. Fujita, E. Takada, T. Endo, and T. Morita. Computerized mass detection in whole breast ultrasound. *Medical Imaging*, 6514, 2007.
- [11] A. Katouzian, E. Angelini, S. Carlier, J. Suri, N. Navab, and A. Laine. A state-of-the-art review on segmentation algorithms in intravascular ultrasound (ivus) images. *IEEE Transactions on Information Technology in Biomedicine*, 16(5):823–834, 2012.
- [12] L. Kuncheva. *Combining Pattern Classifiers*. Wiley-Interscience, Hoboken, New Jersey, 2004.
- [13] F. Lauera, C. Suen, and G. Bloch. A trainable feature extractor for handwritten digit recognition. *Pattern Recognition*, 40(6):1816–1824, 2007.
- [14] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Handwritten digit recognition with a back-propagation network. *Advances in Neural Information Processing Systems*, pages 396–404, 1990.
- [15] Y. LeCun, L. Bottou, Y. Bengio, and H. P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [16] B. Liu, H. Cheng, J. Huang, J. Tian, J. Liu, and X. Tang. Automated segmentation of ultrasonic breast lesions using statistical texture classification and active contour based on probability distance. *Ultrasound in Medicine & Biology*, 35(8):1309–1324, 2009.

- [17] A. Madabhushi and D. Metaxas. Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions. *IEEE Transactions on Medical Imaging*, 22(2):155–169, 2003.
- [18] M. Mancas, B. Gosselin, and B. Macq. Segmentation using a region growing thresholding. *4th Image Processing: Algorithms and Systems*, 56:388–398, 2005.
- [19] Y. Meyer. *Wavelets and Operators*. Cambridge University Press, Cambridge, 1993.
- [20] W. Moon, C. Lo, J. Chang, C. Huang, J. Chen, and C. R. Computer-aided classification of breast masses using speckle features of automated breast ultrasound images. *Medical Physics*, 39, 2012.
- [21] W. Moon, Y. Shen, M. Bae, C. Huang, and J. Chen. Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images. *Pattern Recognition*, 32(7):1191–1200, 2013.
- [22] F. Sahba, M. Tizhoosh, and M. Salma. Segmentation of prostate boundaries using regional contrast enhancement. *IEEE International Conference on Image Processing (ICIP)*, 2:1266–1269, 2005.
- [23] J. Shan, H. Cheng, and Y. Wang. A novel segmentation method for breast ultrasound images based on neutrosophic l-means clustering. *Medical Physics*, 39(9):5669–5682, 2012.
- [24] A. Sohail, P. Bhattacharya, S. Mudur, and S. Krishnamurthy. Classification of ultrasound medical images using distance based feature selection and fuzzy-svm. *Pattern Recognition and Image Analysis*, 6669:176–183, 2011.
- [25] U. Techavipoo, Q. Chen, T. Varghese, A. Zagzebski, and E. Madsen. Noise reduction using spatial-angular compounding for elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 51(5):510 – 520, 2004.
- [26] G. Treece, A. Gee, and R. Prager. Ultrasound compounding with automatic attenuation compensation using paired angle scans. *Ultrasound in Medicine and Biology*, 33(4):630–642, 2007.
- [27] H. Tu, J. Zagzebski, A. Gerig, Q. Chen, E. Madsen, and T. Hall. Optimization of angular and frequency compounding in ultrasonic attenuation estimation. *Journal of the Acoustical Society of America*, 117(5):3307–3318, 2005.
- [28] V. Ulagamuthalvi and D. Sridharan. Automatic identification of ultrasound liver cancer tumor using support vector machine. *International Conference on Emerging Trends in Computer and Electronics Engineering*, pages 41–43, 2012.
- [29] H. Yang, C. Chang, S. Huang, and P. Li. Correlations among acoustic, texture and morphological features for breast ultrasound cad. *Ultrasound Imaging*, 30(4):228–236, 2008.
- [30] M. Yap. A novel algorithm for initial lesion detection in ultrasound breast images. *Journal of Applied Clinical Medical Physics*, 9(4), 2008.
- [31] L. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.
- [32] L. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information Science*, 8:199–249, 1975.
- [33] M. Zhang. *Novel Approaches to Image Segmentation Based on Neutrosophic Logic*. PhD thesis, Utah State University, 2010.