International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 7 (2015) pp. 066-073

Fuzzy Soft Set based Classification for Mammogram Images

Saima Anwar Lashari*, Rosziati Ibrahim and Norhalina Senan

Faculty of Computer Science and Information Technology
Universiti Tun Hussein Onn Malaysia,
86400 Parit Raja, Batu Pahat, Johor, Malaysia
hi120040@siswa.uthm.edu.my,
rosziati@uthm.edu.my, halina@uthm.edu.my

***Abstract*: Mammogram images classification using data mining methods review on past literature showed that these methods are relatively successful however accuracy and efficiency are still outstanding issues. Therefore, the positive reviews produced from past works on fuzzy soft set based classification have resulted in an attempt to use similarity approach on fuzzy soft set for mammogram images classification. Thus, the proposed methodology involved five steps that are data collection, images de-noising using wavelet hard and soft thresholding, region of interest (ROI) identification, feature extraction (statistical texture features), and classification. Hundred and twelve images (68 benign images and 51 malignant images) were used for experimental set ups. Experimental results show better classification accuracy in the presence/absence of noise in mammogram images.**

***Keywords*:** *Mammogram images, feature extraction, wavelet filters, fuzzy soft set, similarity approach on fuzzy soft set*

I. Introduction

Noise present in the digital mammogram images directly influences the competence of a classification task. Study reveals that overall accuracy of classification systems decrease significantly with increase in noise and decrease can become as significant as 21%. This noise can be added due to several factors such as data acquisition and image preprocessing stage. Therefore, the main purpose of de-noising is to remove noisy components (low contrast resolutions, film noise, artifacts) while preserving the important signal as much as possible. De-noising is often done before the images are to be analyzed. Consequently, de-noising plays a very important role in image segmentation, feature extraction and in the classification task [1].

In general, noise filters can be divided into two types, linear methods and non-linear filtering methods. Linear methods involved methods like median, weiner, adaptive and mean filters. Non–linear filtering methods include wavelet transform, multiwavelets and curvelet transform.

Al Jumah [2] worked with various non-linear thresholding techniques such as hard, soft, universal, modified universal and multivariate thresholding in multiwavelet transform domain such as Discrete Multiwavelet Transform Symmetric Asymmetric (SA4), Chui Lian (CL), and Bi-Hermite Bih52S) for different Multiwavelets at different levels to denoised an image and determine the best one out of it. It is found that CL Multiwavelet transform in combination with modified universal thresholding has given best results.

Moreover, in the study of Ramani *et al.* [3], different filters such as median, weiner, adaptive median and mean filters were used in order to remove noise from mammogram images. From their reported results, adaptive median filter performed well then the others filter. The mean square error (MSE) value was small for adaptive median filter 8.4131(mdb 001) and peak signal-to-noise ratio (PSNR) was reported as 38.8812 db (mdb 001).

Saha *et al.* [4] de-noised mammogram images by investigating the role of the embedded thresholding algorithm. Wavelet and curvelet transform were used using soft, hard and block thresholding techniques and compared the performance of the thresholding techniques along with the transforms. Hard threshold using either the transforms performs better over all other techniques from poisson noise removal of mammograms.

So far, few wavelet filters like daubechies db3 and haar have been used for de-noising of medical images [5]. Sidh *et al*. [5] summarizes two wavelet filters namely harr and db3 with additive speckle noise. Their findings suggested that db3 wavelet is more efficient as compare to haar filter. De-noising was performed with different medical images such as MRI, ultrasound, CT-scan and x-ray images. The imposed noise was speckle noise with noise level σ =0.1. The best PSNR value occurs with dataset MRI images with 39.1906 db (hard threshold) and 40.5521 db with soft threshold. Likewise, Malar et al. [6] de-noised mammogram images with three mathematical transform namely wavelet, curvelet and contourlet by hard thresholding.

On the other hand, in our previous work [7], we applied hard and soft threshold functions (using universal threshold function) with different wavelet filters to reduce noise in mammogram images. The purpose of study was to observe the viability of wavelet filters for mammogram images. The highest PSNR for db3 filter was 48.7914 db (hard threshold) and 47.89294 for soft threshold functions. The experimental results are also helpful to select the best wavelet transform for the de-noising of particular medical images such as mammogram images. Thus, the paper provides an alternative de-noising filter for mammogram images. Moreover, it was found that hard threshold is more suitable for mammogram images since images edges were kept and noise was almost suppressed.

In one of the attempts, soft set theory has shown to be capable of handling problems related to classification in particular texture classification, musical instrument classification and decision making problems. For example, Mushrif and Ray [8], presented a novel method for classification of natural textures using the notions of soft set theory, all features on the natural textures consist of a numeric (real) data type, have a value between $[0,1]$ and the algorithm used to classify the natural texture is very similar to the algorithm used by Roy & Maji [9] in the decision making problems.

Senan *et al.* [10] reported the applicability of soft set theory for feature selection of traditional Malay musical instrument sounds. Subsequently, in our previous work [11] performance of two selected classification algorithms based on fuzzy soft set for classification for medical data (numerical data) were evaluated. The acquired results shows that both approaches based on fuzzy soft set performed well, obtaining a classification accuracy reaching 90% for both classification algorithms. Moreover, the experiments conducted demonstrated the effectiveness of fuzzy soft set for medical data categorization.

However, to the best of our knowledge, similarity approach on fuzzy soft set is not yet applied in mammogram images. Therefore, to conduct this study, the proposed methodology involved five steps that are data collection, images de-noising using wavelet hard and soft thresholding, region of interest (ROI) identification, feature extraction (statistical texture features), and classification.

Hundred and twelve images (68 benign images and 51 malignant images) were used for experimental set ups.

The rest of the paper as follows: wavelet thresholding de-nosing, hard and soft threshold functions are presented in Section II. The proposed methodology is given in Section III, experimental results and discussion is reported in Section IV respectively. Finally, the overall conclusions of this study are presented in Section V.

II. Wavelet Thresholding De-noising

Wavelet is a flexible tool which has enormous potential in many applications such as data compression, fingerprint encoding and in image processing. Wavelet-based techniques have raise interest amongst the medical community. Moreover, wavelet techniques have also widely been used in noise reduction, detection of microcalcifications, image analysis and image enhancement.

Wavelet thresholding de-noising is based on the underpinning concept of noise energy is distribute in all wavelet coefficients, while the original signal energy is found in some of the coefficients. Therefore, the signal energy is found much larger than noise energy. So, small coefficients can be considered as caused by noise while large coefficients are triggered by significant signal features [12].

Wavelet threshold de-noising is a very efficient method, the purpose of which is to remove identically distributed gaussian noise [12-13].

Let $x(t) = \{x1(t), x2(t)......... , xn(t)\}$ be the signal series acquired by means of a senor. This signal series consists of impulses and noise. $x(t)$ can be expressed as follows [13].

$$x(t) = p(t) + n(t) \qquad (1)$$

Where

$p(t) = \{p1(t), p2(t)......... , pn(t)\}$ indicates identically distributed and in depended Gaussian noise with mean zero and standard deviation σ . The wavelet threshold de-nosing producer has following steps [13]

1. Transform signal $x(t)$ to the time-scale plane by means of a wavelet transform. It is possible to acquire the results of the wavelet coefficients on different scales.
2. Assess the threshold λ and in accordance with the establish rules , shrink the wavelets coefficients
3. Use the shrunken coefficients to carry out the inverse wavelet transform. The series recovers is the estimation of impulse $p(t)$

The second step has a great impact upon the effectiveness of the procedure. According to Donoho [12], the universal threshold rule should be applied in the second step. According to him, the universal threshold is defined as follows [114][15].

$$\lambda = \sigma\sqrt{2InN} \qquad (2)$$

where

σ refers to the standard deviation of the noise

whereas,

N refers to the number of data samples in the measured signal.

A. *Thresholding*

Thresholding is one of important steps to remove noise. Thresholding is a process of shrinking the small absolute coefficients value while retaining the large absolute coefficient value. It will produce finer reconstructed signal. Since this method is taking the condition that the amplitude of wavelet transform coefficients signals are much larger than noises, so the unconsidered noise will be removed while holding the significant signal[7]. Thus, thresholding is used to segment an image by setting all pixels whose intensity values are above threshold to a foreground value and all the remaining pixels to a background value [14].

Donoho [15] is the person who first introducing the word ‘de-noising’ to explain the process of noise reduction in threshold. Thresholding is mainly divided into two categories: hard thresholding and soft thresholding.

B. Hard Thresholding

The hard-thresholding function used by Donoho is [14-15]

$$\tilde{w}_{j,k} = \begin{cases} \tilde{w}_{j,k} & \left|w_{j,k}\right| \geq \lambda \\ 0 & \left|w_{j,k}\right| < \lambda \end{cases} \tag{3}$$

It is called keep or kill, keep the elements whose absolute value is greater than the threshold. Set the elements lower than the threshold to zero, where $\tilde{w}_{j,k}$ the signal is, λ is the threshold.

C. Soft Thresholding

The soft thresholding function used by Donoho is [14-15]

$$\tilde{w}_{j,k} = \begin{cases} \operatorname{sgn}(w_{j,k})(\left|w_{j,k}\right| - \lambda) & \left|w_{j,k}\right| \geq \lambda \\ 0 & \left|w_{j,k}\right| \geq \lambda \end{cases} \tag{4}$$

It is called shrink or kill which is an extension of hard thresholding, first setting the elements whose absolute values are lower than the threshold to zero and then shrinking the other coefficients where $\operatorname{sgn}(*)$ is symbol function:

$$\operatorname{sgn}(n) = \begin{cases} 1 & n > 0 \\ -1 & n < 0 \end{cases} \tag{5}$$

Peak signal to noise ratio (PSNR) values can be calculated by comparing two images one is original image and other is distorted image. The PSNR has been computed using the following formula [16];

$$PSNR = 10 \log_{10}\left(\frac{R^2}{MSE}\right) db \tag{6}$$

where

R is the maximum fluctuation in the input image data. For example, if the input image has a double precision floating point data type, then R is1. If it has an 8-bit unsigned integer data type, R is 255, etc.

III. Modeling Process for Mammogram Image Classification

The proposed methodology involved five steps that are data collection, images de-noising using wavelet hard and soft thresholding, region of interest (ROI) identification, feature extraction (statistical texture features), and classification as shown in Figure 1. The data was collection from the Mammographic Image Analysis Society (MIAS). There are 322 images, which belong to three categories: normal, benign and malign. These images comprise of 208 normal images, 63 benign and 51 malign. Hundred and twelve images (68 benign images and 51 malignant images) were used for experimental set ups.

For de-noising mammogram images five wavelet filters namely sym8, coif1, haar, db3and db4 at certain level of hard and soft threshold functions have been used. Different wavelet filters with obtained PSNR values for hard and soft threshold functions have been calculated and reported.

The purpose of identifying region of interest (ROI) is to encompass exclusively on the appropriate breast region which reduces the possibility for erroneous classification. Later on, feature extraction has been done. Feature extraction is a process of computing numerical representation of the mammogram images. It is believed that strong feature set will likely provides better classification accuracy rate. Thus, six statistical features which are mean, variance, skewness, kurtosis, contrast and smoothness are extracted from ROI of the mammogram images [17].

Once a set of baseline feature measurements was made, data normalization is done. The transformed data in the set of interval $[0,1]$ are more cognitively relevant for human understanding and computation process goes faster. For classification, an algorithm based on similarity approach on fuzzy soft set has been applied in mammogram images.

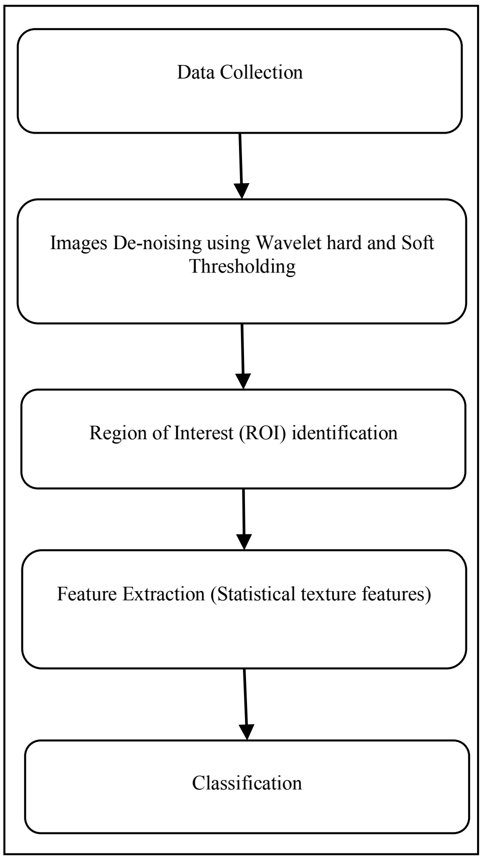


Figure 1. Modeling Process for mammogram image classification

Six statistical features which are mean ,variance, skewness, kurtosis, cotrast and smoothnes are computed from ROI. The mathematical model to compute these featuer set is listed in Table 1.

Sr. no	Feature	Formula	Description
1.	Mean	$\sum_{i=0}^{L-1} z_i * p(z_i)$	A measure of average intensity
2.	Variance	$\sum_{i=0}^{L-1} (z_i - m)^2 * p(z_i)$	Second moment about the mean
3.	Skewness	$\sum_{i=0}^{L-1} (z_i - m)^3 * p(z_i)$	Third moment about the mean
4.	Kurtosis	$\sum_{i=0}^{L-1} (z_i - m)^4 * p(z_i)$	Fourth moment about the mean
5.	Contrast	$\sum_{i=0}^{L-1} \sqrt{(z_i - m)^2 * p(z_i)}$	Standard deviation of pixel intensities
6.	Smoothness	$1 - \frac{1}{(1+\sigma^2)}$	Measures the relative intensity variations in a region

Table 1. Satistical features

Where

$z_i \rightarrow$ is a random variable indicating intensity

$p(z_i) \rightarrow$ is the histogram of the intensity levels in a region

$L \rightarrow$ is a number of possible intensity levels

$\sigma \rightarrow$ is the standard deviation

A. Similarity Approach on Fuzzy Soft set based classification

As described in the work of [18], this classifier uses similarity between two fuzzy soft sets to classify numerical data. In this paper, to classify mammogram images with similarity approach on fuzzy soft set classifier, by adding the second step in both training and classification phase i-e the second step is wavelet de-nosing with normalization process (refer [18] having similar algorithm with [8]). Fuzzy soft based classifier learns by calculating the average value of each parameter (attribute or feature) from all object or instant with the same class label, to construct fuzzy soft set model with universe consisting of all of class labels. In other words, an object in the universe represents all data derived from the same class label. Furthermore, to classify the test data with unknown class labels, first, construct the fuzzy soft set model of data using formula as stated in equation 7, and then find similarity between two fuzzy soft sets as described in the work of Majumdar & Samanta [19] as stated in equation 9). The next is to calculate the class similarity score to determine the class label as stated in equation 10.

Feature fuzzification can be done by dividing each attributes value with the largest value at each attributes [18].

$$e_{fi} = \frac{e_i}{\max(e_i)} \qquad (7)$$

where $e_i, i = 1,2,....., n$ is the old attribute and e_{fi} is attribute with new value between $[0,1]$.

B. Classification Algorithm

Classification algorithm is divided into two phases namely training phase and classification phase. The description on both phases is given below:

Training phase

1. Given N samples obtained from the data class w.
2. De-noised images using wavelet hard and soft threshold functions using equation 3, 4 and 5 and obtain a feature vector E_{wi}, for i=1,2, …, N.
3. Feature fuzzification to obtain a feature vector E_{wi}, for i=1,2, …, N for all training data using equation 7
4. Calculate the cluster center vector E_w $i = 1,2,.., N$.

$$E_w = \frac{1}{N}\sum_{i=1}^{N} E_{wi} \qquad (8)$$

5. Obtain fuzzy soft set model $(\tilde{F}, E)$ for class w is a cluster vector for class w having D features
6. Repeat the process for all W classes.

Classification phase

1. Obtain the unknown class data.
2. De-noised images using hard and soft threshold functions using equations 3, 4 and 5 and obtain a feature vector E_{wi}, for i=1,2, …, N
3. Feature fuzzification to obtain a feature vector E_{wi}, for i=1,2, …, N for all testing data using equation 7
4. Obtain a fuzzy soft set model for unknown class data $(\tilde{G}, E)$
5. Compute similarity between $(\tilde{G}, E)$ and $(\tilde{F}, E)$ for each w using equation 9

$$S(F_\rho, G_\delta) = M_i(\tilde{F}, \tilde{G}) = 1 - \frac{\sum_{j=1}^{n} \left|\tilde{F}_{ij} - \tilde{G}_{ij}\right|}{\sum_{j=1}^{n} \left(\tilde{F}_{ij} + \tilde{G}_{ij}\right)} \qquad (9)$$

6. Assign the unknown data to class w if similarity is maximum $w = \arg\left[\max_{w=1}^{W} S(\tilde{F}, \tilde{G})\right]$ (10)

where

E_w is mean of feature vectors in the same class label.

IV. Results and Discussion

Preprocessing helps to enhance the contents mammogram images and removal of noise from images will increase the quality of images and afterward contributes in increase of classification accuracy rate. Table 2 shows five types of filters namely sym8, coif1, haar, db3and db4 at certain level of hard and soft threshold functions which are used for de-noising mammogram images mainly concentrating on the peak signal to noise. Different PSNR values are calculated at different levels of gaussian noise at hard and soft thresholding levels by applying these filters one after the other and then comparison is made.

The best PSNR value is 46.44656db (hard threshold) and 43.80779 db (soft threshold) which occurs in db3 wavelet filter. On other hand, mammogram images have good PSNR values; it could be these images have high fine details edge that is the reason that hard thresholding produce better PSNR values than soft thresholding.

Types of filters	Types of threshold	Noise Levels	PSNR(db)
Sym8	Hard	σ =10	46.01435
		σ =20	42.74203
		σ =40	40.98273
	Soft	σ =10	43.4963
		σ =20	41.63135
		σ =40	40.98034
Db3	Hard	σ =10	46.44656
		σ =20	42.53113
		σ =40	41.64176
	Soft	σ =10	43.80779
		σ =20	41.84176
		σ =40	41.64176
Db4	Hard	σ =10	45.30231
		σ =20	42.76131
		σ =40	42.32381
	Soft	σ =10	43.67878
		σ =20	42.49307
		σ =40	42.32194

Table 2. PSNR values of MIAS after processing through different wavelet filters

Types of filters	Types of threshold	Noise Levels	PSNR(db)
haar	Hard	σ =10	45.91019
		σ =20	43.76875
		σ =40	39.53933
	Soft	σ =10	43.25596
		σ =20	40.68419
		σ =40	39.0267
Coif1	Hard	σ =10	45.97968
		σ =20	42.05187
		σ =40	40.41731
	Soft	σ =10	43.12949
		σ =20	40.88994
		σ =40	40.41376

Table2. PSNR values of MIAS after processing through different wavelet filters (Cont.)

Table 3 depicts different filters with best obtained PSNR values for hard and soft threshold functions. From the observations, db3 provides better results while compared with the other filters for purpose of de-noising for mammogram images. The best PSNR value is 46.44656db (hard thresholding) and 43.80779 db (soft thresholding). To sum up pre-processing through five wavelet filters, db3 wavelet filter with noise level σ =10 is more suitable for mammogram images.

	Types of threshold	Filter Sym8	Filter Db3	Filter Db4	Filter haar	Filter Coif1
Mammogram Images	Hard	46.014	46.44	45.302	45.910	45.980
	Soft	43.49	43.80	43.679	43.255	43.130

Table 3. PSNR values for MIAS after processing through different wavelet filters

Effectiveness of the proposed classification algorithm for mammogram images have been thoroughly tested using hundred and twelve images (68 benign images and 51 malignant images). MIAS dataset has been divided into parts: 70% for training and 30% testing and obtained features have been normalized to form a fuzzy value between $[0,1]$.

For different experimental setups, at least 10 times, train and test data were selected randomly. From Table 4, it can observe that soft threshold provides better classification rate than hard threshold. The highest classification rate occurs with filter db3 (Level 4) with accuracy 62.12 % (soft threshold) with cpu time 0.0026sec.

Wavelet de-nosing filters with different decomposition levels			Fuzzy soft set Classifier
		Accuracy (%)	CPU Time (Sec)
Daubechies db3 (Level 1)	Hard Thresholding	55.15	0.0017
	Soft Thresholding	58.18	0.0027
Daubechies db3 (Level 4)	Hard Thresholding	54.55	0.0027
	Soft Thresholding	62.12	0.0026
Sym8 (Level 1)	Hard Thresholding	56.06	0.0022
	Soft Thresholding	50.00	0.0029
Sym8 (Level 4)	Hard Thresholding	59.09	0.0038
	Soft Thresholding	57.58	0.0028

Table 4. Classification accuracy and CPU time for different wavelet filters

For the comparison purpose from literature, the work of Naveed *et al.* [1] has been chosen since their work reported classification accuracy with and without noise. Different classifiers like k-nearest neighbor (KNN), Bayesian, support vector machine and neural network (NN) were used. Quantum and impulse noise filtering were used to observed accuracy with and without noise. It was reported that noise presence badly affect the classification accuracy.

Table 5 shows the comparison of classification accuracy in the presence/absence of noise in mammogram images. It is observed that presence of noise badly affect the classification accuracy. The proposed method with de-noising filter provides accuracy 62.12% (with noise) and 63.64% (without noise) which is comparatively better than other reported techniques such as NN and Bayesian. However, the classifiers SVM and K-NN have better accuracy without noise. Moreover, the six statistical features which are mean, variance, skewness, kurtosis, contrast and smoothness are strong feature set. These feature set help in producing better classification results with and without noise.

Technique	**Mammogram with noise accuracy (%)**	**Mammogram with accuracy (%) without noise**
NN+ features [1]	58.2 (Poisson noise) 56.3 (Salt and Pepper noise)	63.6
Bayesian+ features [1]	59.1 (Poisson noise) 57.5 (Salt and Pepper noise)	63.1
Proposed method	62.12 (Gaussian noise)	63.64

Table 5. Classification accuracy with and without noise

V. Conclusion

In this paper, an algorithm based on similarity approach on fuzzy soft set with noise removal filter is presented. Results were evaluated and compared with different leading noise filtration techniques from recent literature. The empirical evidences states that proposed classification algorithm performs better than the existing classification algorithm with different noise types (poison noise and salt and pepper noise). Moreover, different PSNR values are calculated with gaussian noise on mammogram images at different level of hard and soft threshold functions by applying these wavelets filters techniques namely haar, db3, coif1, sym8 and db4 filters one after the other and then comparison is made.

The obtained results shows the superiority of proposed classification algorithm in terms of noise filter based on performance measure accuracy and cpu time. These experiments are also helpful to select the best wavelet transform for the de-noising of particular medical images such as mammogram images Thus; this paper provides an alternative de-nosing method for mammogram images. Moreover, this study contributes by extending the robustness of fuzzy soft theory into examining mammogram images within medical image classification domain. To the best of our knowledge, this theory is mainly used within texture classification, and decision making problems context.

Acknowledgment

The authors would like to thank office for Research, Innovation, Commercialization and Consultancy Management (ORICC) and Universiti Tun Hussein Onn Malaysia for supporting this research under vote no U110.

References

[1] N.Naveed, A. Hussain, M. A. Jaffar, and T. S. Choi, "Quantum and impulse noise filtering from breast mammogram images, "*Computer methods and programs in biomedicine*, *108*(3), 1062-1069, 2012.

[2] A. Al Jumah, M.. G. Ahamad and S. A. Ali., " Denoising of Medical Images Using Multiwavelet Transforms and Various Thresholding Techniques," Journal of Signal and Information Processing 4 pp. 24-32, 2013.

[3] R. Ramani, N. S. Vanitha and S.Valarmathy, "The Pre-Processing Techniques for Breast Cancer Detection in Mammography Images," *International Journal of Image, Graphics & Signal Processing*, *5*(4), 2013.

[4] M. Saha, M. K. Naskar, and B. N, Chatterji, "Soft, Hard and Block Thresholding Techniques for Denoising of Mammogram Images," *IETE Journal of Research*, (ahead-of-print), 1-6, 2015.

[5] K. Sidh, B. Khaira, and I. Virk, " Medical image denoising in the wavelet domain using haar and DB3 filtering," *International Refereed Journal of Engineering and Science*, *1*, 1-8, 2012

[6] E. Malar, A.Kandaswamy, S. S. Kirthana, D. Nivedhitha, and M. Gauthaam,. "Curvelet image denoising of mammogram images,". *International Journal of Medical Engineering and Informatics*, *5*(1), 60 -67, 2013.

[7] S. A. Lashari, R. Ibrahim and N. Senan, "De-noising Analysis of Mammogram Images in the Wavelet Domain using Hard and Soft Thresholding," Proceeding of Fourth World Congress on Information and

Communication Technologies (WICT 2014), December 08-11, 2014, Malacca, Malaysia

[8] M. M. Mushrif, S. Sengupta, and A. K. Ray, "Texture classification using a novel soft set theory based classification algorithm, " In: LNCS, Volume 3851, pp.246-254, Springer, Heidelberg, 2006.

[9] A. R. Roy, and P. K Maji, "A fuzzy soft set theoretic approach to decision making problems," Journal of Computational and Applied Mathematics, 203 (2), 412-418, 2007.

[10] N. Senan, R. Ibrahim, N. M. Nawi, I. T. R. Yanto and T. Herawan "Soft set theory for feature selection of traditional Malay musical instrument sounds", In *Information Computing and Applications* (pp. 253-260). Springer Berlin Heidelber, 2010.

[11] S. A. Lashari, and R. Ibrahim, "Performance Comparison of Selected Classification Algorithms Based on Fuzzy Soft Set for Medical Data," In *Advanced Computer and Communication Engineering Technology* (pp. 813-820). Springer International Publishing, 2015.

[12] I: Daubechies, "*Ten lectures on wavelets",* Philadelphia: Society for industrial and applied mathematics, Vol. 61, pp. 198-202, 1992.

[13] Z. Zang, Wang and Y. Zheng, "Analysis of signal de-noising method based on an improved wavelet thresholding," In *Electronic Measurement & Instruments, 2009. ICEMI'09. 9th International Conference on* pp. 1-987, IEEE August 2009.

[14] D. L. Donoho, "De-noising by soft-thresholding," *Information Theory, IEEE Transactions on, 41*(3), 613-627, 1995.

[15] D. L. Donoho, I. M . Johnstone, G. Kerkyacharian, and D. Picard, "Wavelet shrinkage: asymptopia?," *Journal of the Royal Statistical Society. Series B (Methodological)*, 301-369, 1995.

[16] K. Sidh, B. Khaira and I. Virk, Medical image denoising in the wavelet domain using haar and DB3 filtering, *International Refereed Journal of Engineering and Science*, *1*, 1-8, 2012.

[17] R. Aarthi, K. Divya, and S. Kavitha, "Application of Feature Extraction and Clustering in Mammogram Classification using Support Vector Machine," Third International Conference on Advanced Computing (ICoAC), pages 62-67, 2011.

[18] B. Handaga, and M. M. Deris**,** "Similarity approach on fuzzy soft set based numerical data classification, "Communications in Computer and Information Science. Volume 180(6), pp. 575-589, 2011.

[19] P. Majumdar,. and S. K Samantra, "Generalized fuzzy soft set" Journal of Computational and Applied Mathematics. Pp. 1279-1286 2010.

Author Biographies

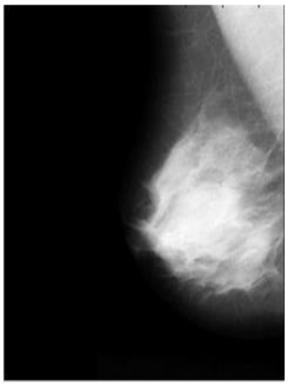
Saima Anwar Lashari was admitted to the PhD degree in the Faculty of Computer Science and Information Technology at the Universiti Tun Hussein Onn Malaysia (UTHM), Malaysia in 2013.She received her MSc degree in Information Technology in 2012 at Universiti Tun Hussein Onn Malaysia (UTHM), Malaysia and her Bachelors (Hons) in Computer Science in 2004 at University Of Engineering & Technology (UET), Pakistan. Her research interest includes data mining, classification, soft set.

Rosziati Ibrahim obtained her Bachelor of Science (Hons) in Applied Mathematics from University Of Adelaide in 1989. She holds an Msc degree in Applied Mathematics from the same university in 1991. She completed her PhD in Information Technology (Software Specification) at Queensland University of Technology (QUT) in 2000. Her research interest includes software engineering, image processing and machine learning.

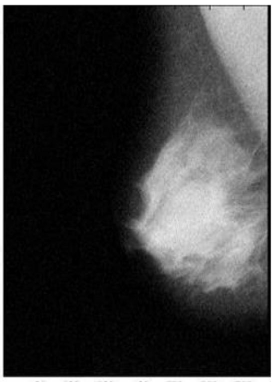
Norhalina Senan obtained her Bachelor of Science in 2000 and Msc degree in 2004 from Universiti Teknologi Malaysia. She completed her PhD in Information Technology (Soft computing) at Universiti Tun Hussein Onn Malaysia (UTHM) in 2013.Her research interest includes data mining, classification, rough set, feature selection, signal processing, multimedia content development.

Appendix A: Results with Visu Shirnk

Original l image (mdb001)



Gaussian Noise 10%

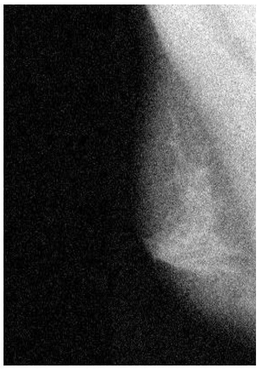


Visu Shirnk by Hard Thresholding Sym8 at level 1

Visu Shirnk by Soft Thresholding Sym8 at level 1

Original l image (mdb213)

Gaussian Noise 20%



Visu Shirnk by Hard Thresholding db3 at level 1

Visu Shirnk by Soft Thresholding db3 at level 1