Journal of College of Education for Pure Sciences
Volume 8, Number 2, June 2018
Website: jceps.utq.edu.iq Email: jceps@eps.utq.edu.iq

Breast Cancer Diagnosis Using Wavelet and Fuzzy Logic

Shaker K. Ali , Wamidh K. Mutlag

shaker@utq.edu.iq , alsalhi55@yahoo.com

(1) University Computer Sciences and Mathematics College, Computer Department, of Thi-Qar, Iraq

(3) Computer Department ,College of Education for Pure Science's,

(4) University of Thi-Qar

Abstract:

Breast cancer (BC) is one of the important general health problem in the world. There are two types of Breast Cancer; benign breast cancer and malignant breast cancer .In this paper we suggest a new algorithm to diagnosis the two types of BC. Our algorithm has three steps; the first step is pre-processing ,then feature extraction by using wavelet transform and the results from the wavelet transform minimized by using standard division this is for second step and the third step is for diagnosis by using Fuzzy logic , the results from our algorithm are 98%. The data base usedfrom http://web.inf.ufpr.br/vr/breast-cancer-database .

Keyword: Breast cancer, diagnosis,fuzzy logic ,wavelets transform

1. Introduction

Cancer is a substantial public medical condition nowadays. Based on the IARC (International Firm for Research on Cancers) of the WHO (World Health Corporation), 8.2 million fatalities were brought on by cancer tumor in 2012 and 27 million of new circumstances of the disease are anticipated before 2030[1].

Proper medical identification can help a person to eliminate BC associated risk if the health of cancers tumor type is safe (localized and non - invading) or malignant (invading and life intimidating). For analysis, different methods like mammograms, CT check, histopathology image research (biopsy image) are usually used. When there's a sure result of mammograms about BC, then prognosis with histopathology imaging can be executed [2].

A histology image research system generally has a mixture of hardware and software and maybe it's put into two consecutive subsystems: (1) composition prep and image creation and then (2)image managing analysis. To lessen the death rate among women two thing are extremely important that are education about breasts cancers and verification that means acknowledgement[2] In This paper we didn't use hardware only software of diagnosis .

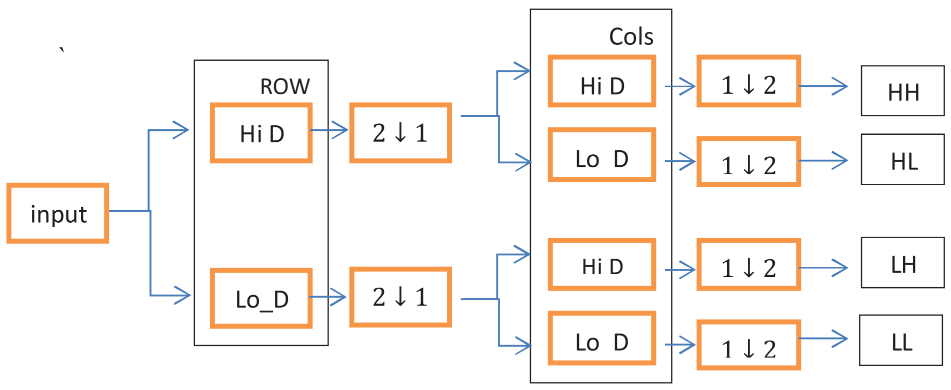
The paper organized as follows: In Section 2, the research methodology is presented. In this section we also present all methods incorporated in the research method. Section 3 presents Algorithm Suggested . Section 4 presents the results . Finally, in Section 5the conclusion and present future work.

2. Methodology

We use wavelet transform for image analysis to find the important features that is the key of diagnosis which used in Fuzzy logic after find the specific feature by using standard deviation for find the feature vector.

2.1. Haar Wavelet Transform (HWT)

Discrete Wavelet Transform (DWT) is one of the evaluation image techniques in change domain. This technique has been created as an extremely efficient and versatile way for decomposing alerts into four sub-bands[3]. DWT decomposes image into four parts using two kinds of filters. The first filter is Low Pass Filter and the second filter is High Pass Filter. Subband image are LL, HL, LH, HH. The first Subband is LL which represents a low-frequency values, the second sub-band is LH presents vertical information and the third sub-band is HL symbolizes horizontal information whereas for diagonal information is represented by HH[4]. According to [5]important values are the lowest sub-band (LL). In its development many reports to improve the performance of DWT filtration systems. Haarmother function hastrusted filtering methods in wavelet transforms. DWT with Haarfiltration system known as HaarTransform (HWT). This filters has advantages such as efficient memory usage, fast and simple [6]. Here is the decomposition process conducted by Haar filter.[6]



Fig(1) Signal Decomposition with Haar filtration systems[6]

The Haar Wavelet transform equation is [7] :

$$\lambda G_0(\omega) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} e^{-i\omega} = \sqrt{2e^{-\frac{i\omega}{2}}} \cos\left(\frac{\omega}{2}\right) \quad 1$$

$$\lambda G_0(\omega) = \frac{1}{\sqrt{2}} e^{i\omega} - \frac{1}{\sqrt{2}} = \sqrt{2ie^{-\frac{i\omega}{2}}} \sin\left(\frac{\omega}{2}\right) \quad 2$$

2.2. Standard Deviation (STD)

After decomposing the images usingWavelet Transform, we can extract some importantfeatures using Mean, Standard Deviation. The mean is implemented as follows[8].

$$mean = \frac{1}{RS} \sum_{r=0}^{R-1} \sum_{s=0}^{S-1} f(r,s) \quad 3$$

The Standard Deviation (STD) is implemented as[8]:

$$STD = \sqrt{\frac{\sum_{r=0}^{R-1} \sum_{s=0}^{S-1} (f(r,s) - mean)^2}{R \times S}} \quad 4$$

2.3. Fuzzy Logic

Fuzzy logic examines human thinking and represent this thinking as mathematical rule to applies it to conduct problem solving and decision making. Verbal rules and variables used in the human decision-making process are fuzzy, unlike the precise, numeric nature of computer logic. These verbal terms are represent mathematically as membership functions. The decision-making in fuzzy logic using symbolic verbal phrases rather than numeric values and can produce best results depend on indefinite verbal knowledge like humans. If a system'sbehavior can be modelled by rules or requires very complex nonlinear processes, fuzzy logic can be applied to this system. Mamdani's fuzzy inference method is the most commonly used fuzzy inference system.

Fuzzy rule-based systems consist of three main steps; Fuzzification, Inference and Defuzzification as illustrated in Figure 2. In the fuzzificationstep, the crisp input and output variables are defined and mapped to linguistics variables. Once the input and output variables and the corresponding membership functions are defined, a system of rules composed of IF-THEN statements is designed and the fuzzy inference takes place, producing a fuzzy output set. The output is then defuzzified during the defuzzification step to produce a crisp output value - the prediction [9]. For our case, Range between (0-5) and (10-16) was used to represent the output which is the probability malignant or benign respectively [9].

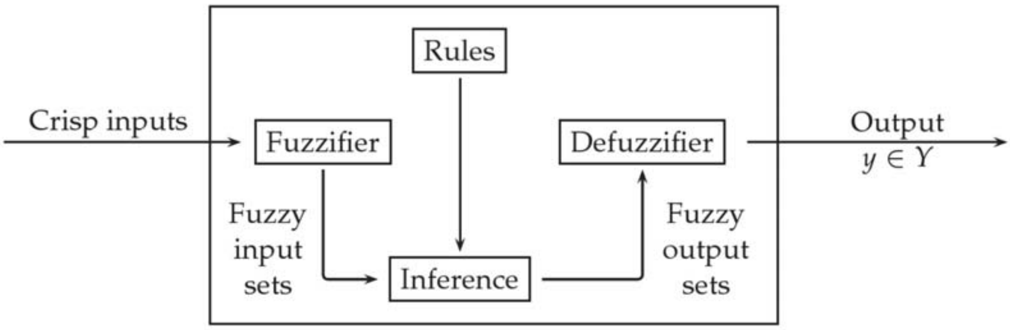


Figure 2 : Fuzzy Inference System[9]

3. The Suggested Algorithm

Our suggested algorithm is consisting of four steps as shown in Figure 3

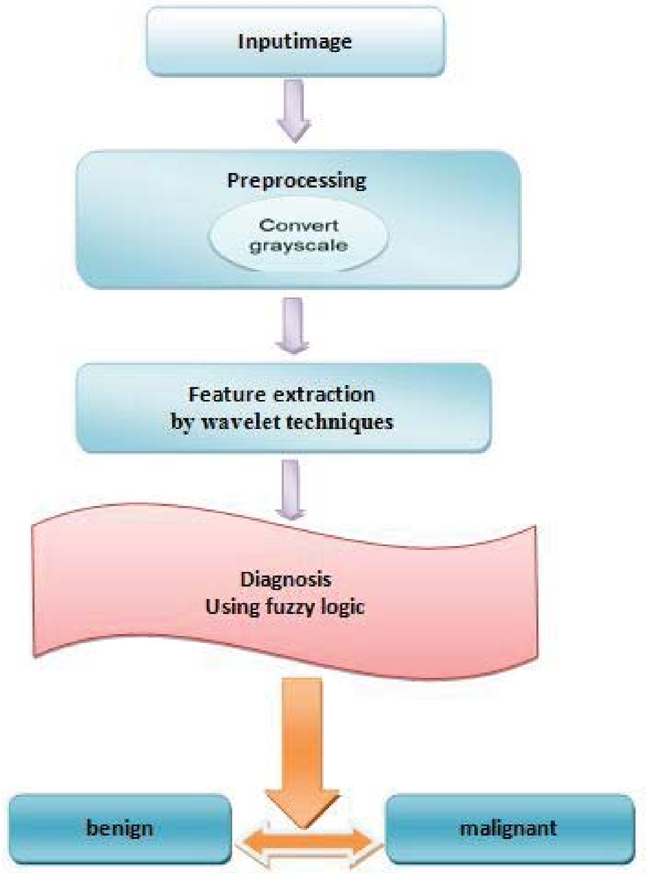
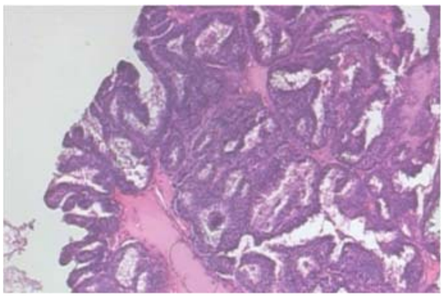


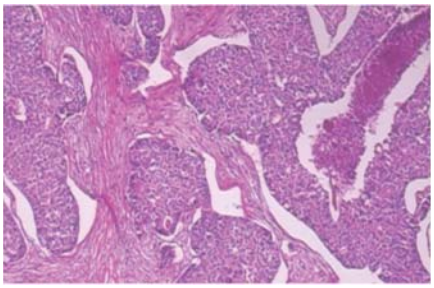
Figure 3 : The suggestedAlgorithm

3.1 image acquisitions:

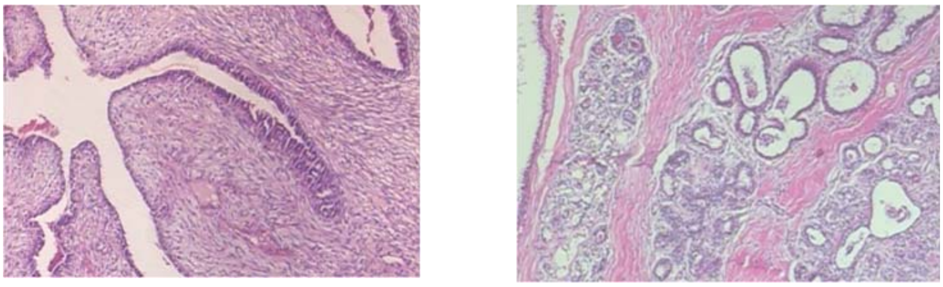
By collect the images database from internet (http://web.inf.ufpr.br/vr/breast-cancer-database) which is contain 80 images for training (40 benign images and 40 malignant images) , and 40 image for testing.as shown in Figure .4 where all image with (460*700) with 4X zoom .



A-: papillary



B-: ductal



C-: phyllodes_tumor D-: adenosis

Figure 4: (A and B) Benign images , (C and D) malignant images

3.2.pre-processing:

1- the pre-processing is starting by transform a histology colour image into greyscale image ; where the color image contain matrix with three color values for each pixel; Red (R), Green (G) and Blue (B), as shown in in Figure 5.

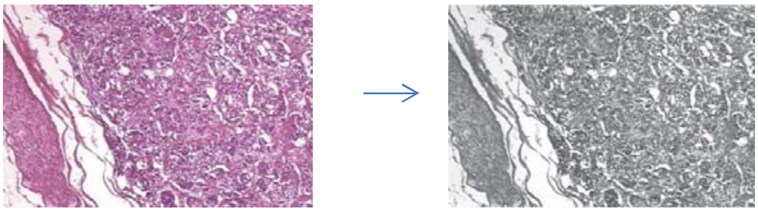


Figure 5:transform a histology image to greyscale image

2- stores all numeric variables as double-precision floating-point values that are 8 bytes (64 bits). These variables have data type (class) double.

3.3. Feature Extraction :

1- Calculate CA1,CA2 and CA3 and CD2 (the approximately parts ofHaar Wavelet Transform(HWT) for three levels respectively);as shown in in Figure.6.

<table>
<tr><td>CA3</td><td>CH3</td><td rowspan="2">CH2</td><td rowspan="3">CH1</td></tr>
<tr><td>CV3</td><td>CD3</td></tr>
<tr><td colspan="2">CV2</td><td>CD2</td></tr>
<tr><td colspan="3">CV1</td><td>CD1</td></tr>
</table>

Figure 6: HWT for three levels

2- . Calculate STD1, STD2 and STD3 and STD4 for CA1, CA2 , CA3 and CD2respectively.There are four features will obtained (STD1, STD 2, STD 3, and STD 4) which be the inputs for the fuzzy logic

3.4.Diagnosis :

1- Fuzzy logic is used for data classification, which is widely used in diseases diagnosisbecause it is efficiency and robustness. We have 4 inputs and one output fuzzy rule.
2- The four input is (STD1, STD2 and STD3 and STD4) from third step, and there are two rules of fuzzy logic we need to diagnosis and one output as shown in Figure.7,and Table 1.

3- Each source has 2 membership functionswhich are corresponding to the patient image histology data worth, and productivity has 2 membership functions which can be matching to types Brest tumor benign or malignant . The inputs including 2 regular membership functions were made to improve accuracy. Figure 7 shows the Mamdani.
4- the model properly recognizes benign and malignant tumors according to the input values, which originated from patient data. Ultimately, the fuzzy inference system is able to predict diagnosis status with better accuracy higher than 98%. A comparison experiment has been conducted as well.As shown in Figure.8.

Table 1: Fuzzy Rules

No	rule
1	IF (input1 is duc+pap) OR (input2 is duc) OR (input3 is duc_+_pap) OR (input4 is duc_+pap) THEN (output1 is B)
2	IF (input1 is ade_+phy) OR (input2 is ade+phy) OR (input3 is ade_+_phy) OR (input4 is ade_+_phy) THEN (output1 is A)

Where; Duc =ductal ,pap =papillary ,ade =adenosis , phy = phyllodes

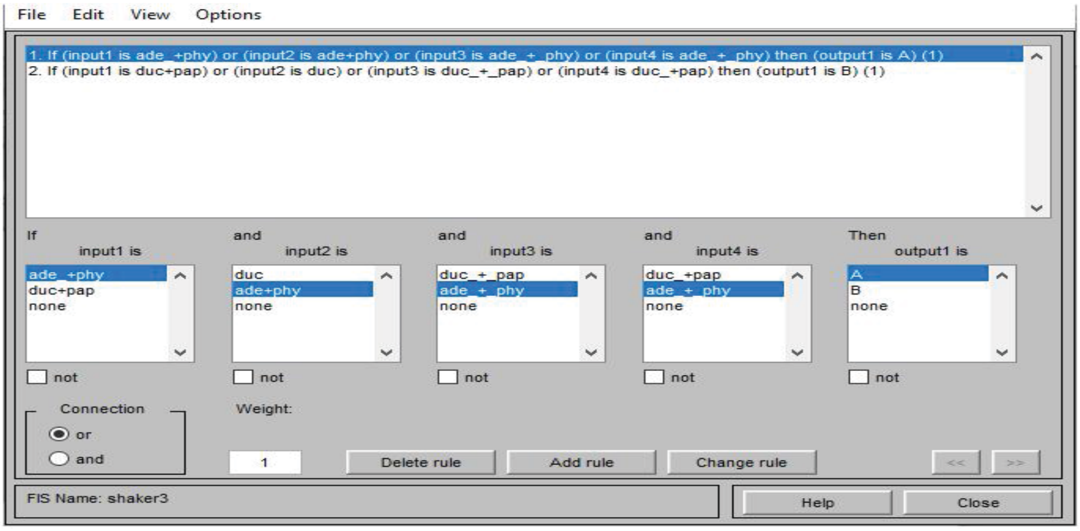
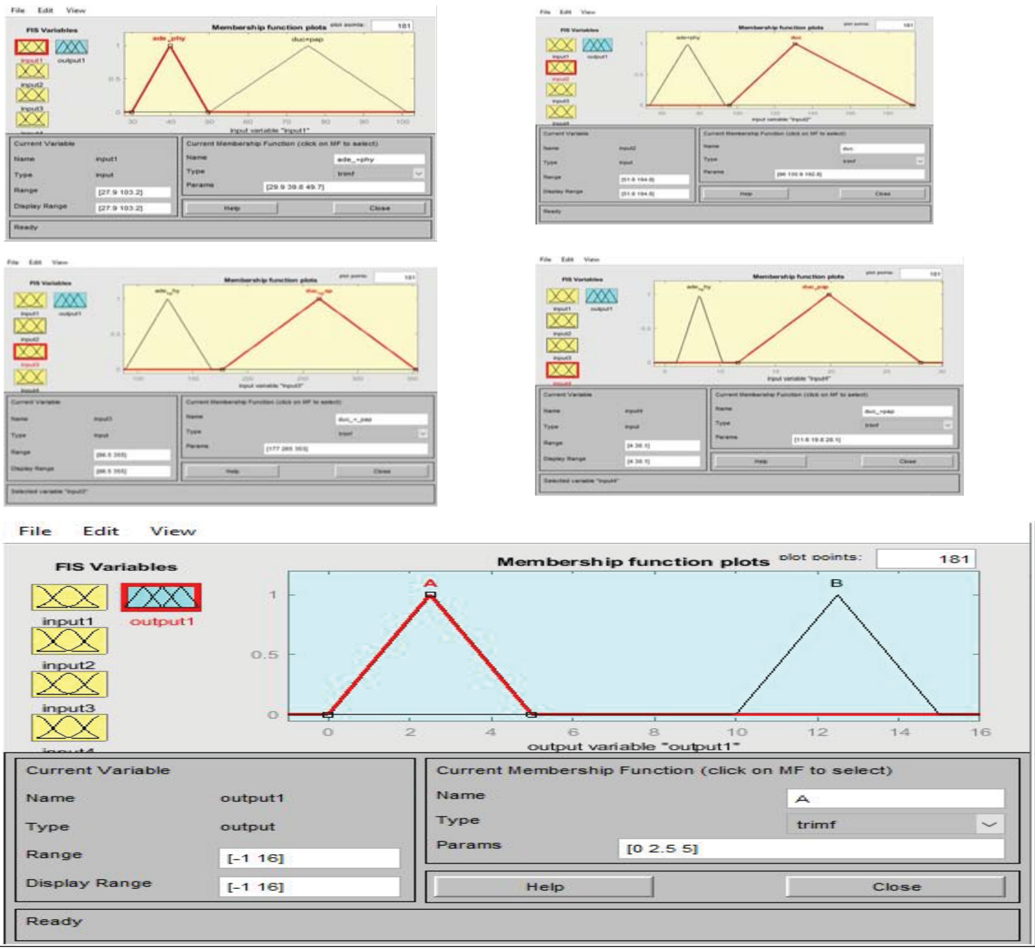


Figure 7:Rule input



4. Results

In this section, we discus several concepts needed for our proposed algorithm . In This paper we use histopathology images for diagnosis the breast cancer through check the images if it is benign (BC) or malignant (BC)by use wavelet algorithm and by extracting the features for each case then these features will be input to fuzzy logic to give the final decision, there are four features used as four input for fuzzy logic which give the final results (benign or malignant) ; the results from our algorithm gives best accuracy and it is benefit for diagnosis are 98% .

There are four features bycalculate STD1, STD2 and STD3 and STD4 for CA1 , CA2 , CA3 and CD2 respectively four features (STD1, STD2, STD3, and STD4) , as shown in table (2 and 3) ,will be the inputs of the fuzzy logic through the approximately parts of HWT for three levels respectively .

Table (2).STD For Benign

Benign				
Image no.	**STD4 CD2**	**STD3 CA3**	**STD2 CA2**	**STD1 CA1**
1	8.366293	144.2333	81.3981	43.63146
2	8.067114	118.6775	69.13245	37.67848
3	6.592461	117.4697	66.36621	35.47462
4	6.059023	128.7396	70.91358	37.28117
5	7.511251	122.6454	70.2219	37.75304
6	8.678011	128.9053	75.18213	40.72568
7	9.10601	129.4749	75.20538	40.92783
8	8.928726	120.7492	71.01208	38.8137
9	9.014804	122.0768	71.32115	38.97185
10	9.09672	121.364	71.96632	39.36198

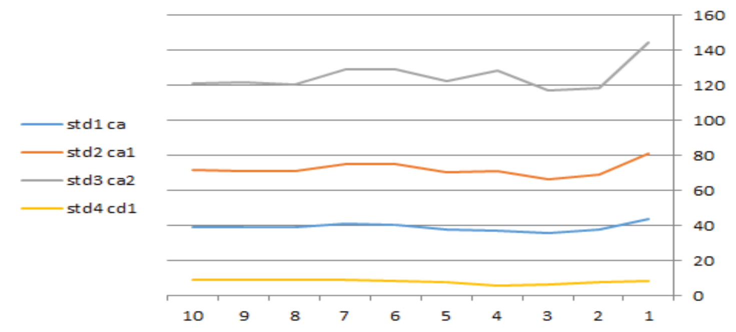


Figure 9 : Benign

Table (3).STD For Malignant

malignant				
Image no.	STD4 CD2	STD3 CA3	STD2 CA2	STD1 CA1
1	22.30879	319.3109	184.3522	**99.56281**
2	24.91326	302.6923	179.1271	**98.53386**
3	28.12394	243.9512	157.2049	**90.35643**
4	24.82369	279.0232	165.576	**91.66428**
5	25.65492	264.8339	164.9793	**92.89925**
6	16.2245	181.6433	107.7856	**59.85532**
7	16.09513	219.1201	124.2312	**67.52448**
8	12.50924	266.9144	141.7642	**73.81036**
9	22.51926	190.222	120.1879	**69.40446**
10	20.23946	315.3208	173.5048	**92.65795**

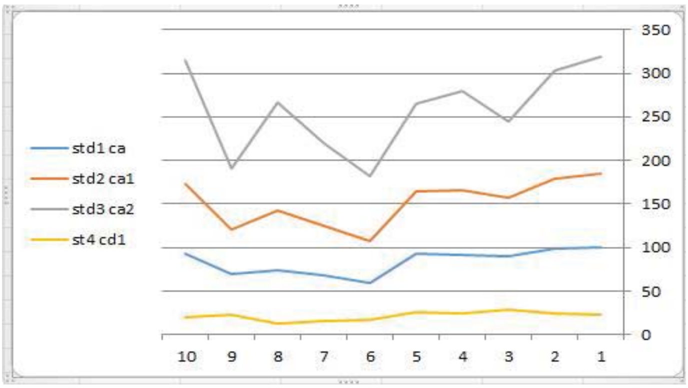


Figure 10 : malignant

5. Conclusion

By using (HWL) into tree levels we can have four features at each level it will be 12 separate vectors CA1 , CH1 , CV1, CD1 , CA2 , CH2 , CV2 , CD2 , CA3 , CH3 , CV3 , CD3) but in the experiments we found that CA1 , CA2 , CA3 and CD2 respectivelyare the best parts will give the best feature for diagnosis .

Many different kinds statistical features are used such as entropy, mean, moment and standard deviation but in the experiments we found that the best ones are STD .

References

1. Spanhol, F.A., et al., *A dataset for breast cancer histopathological image classification*. IEEE Transactions on Biomedical Engineering, 2016. **63**(7): p. 1455-1462.

2. Johra, F.-T. and M.M.H. Shuvo. *Detection of breast cancer from histopathology image and classifying benign and malignant state using fuzzy logic*. in *Electrical Engineering and Information Communication Technology (ICEEICT), 2016 3rd International Conference on*. 2016. IEEE.

3. Gupta, D. and S. Choubey, *Discrete wavelet transform for image processing*. International Journal of Emerging Technology and Advanced Engineering, 2015. **4**(3): p. 598-602.

4. Susanto, A., C.A. Sari, and E.H. Rachmawanto. *Hybrid method using HWT-DCT for image watermarking*. in *Cyber and IT Service Management (CITSM), 2017 5th International Conference on*. 2017. IEEE.

5. Sathik, M.M. and S. Sujatha, *A Novel DWT Based Invisible Watermarking Technique for Digital Images*. Int. Arab J. e-Technol., 2012. **2**(3): p. 167-173.

6. Sudibyo, U., et al., *A Secure Image Watermarking using Chinese Remainder Theorem Based on Haar Wavelet Transform*.

7. Ryan, Ø., G. Dahl, and K. Mørken, *Nonlinear optimization Lecture notes for the course MAT-INF2360*. University of Oslo. 2014.

8. Kaganami, H.G., S.K. Ali, and B. Zou, *Optimal approach for texture analysis and classification based on wavelet transform and neural network*. Journal of Information Hiding and Multimedia Signal Processing, 2011. **2**(1): p. 33-40.

9. Ahadi, F.S., et al. *Feature-Based classification and diagnosis of breast cancer using fuzzy inference system*. in *Information and Automation (ICIA), 2017 IEEE International Conference on*. 2017. IEEE.