Development of Breast Cancer Diagnosis Tool using Hybrid Magnetoacoustic Method and Artificial Neural Network


Abstract—Breast cancer is a metabolic disease that causes the breast cells to acquire genetic alteration and allows them to grow beyond the normal tissue limit. With the yearly increasing trend in new cases and mortality rate, new approach in diagnosis and treatment of breast cancer is crucial to improve the existing management of breast cancer cases. This paper presents a new approach in breast cancer diagnosis by using Hybrid Magnetoacoustic Method (HMM) and artificial neural network. HMM is a newly developed one dimensional imaging system that combines the theory of acoustic and magnetism for breast imaging. It is capable to produce 2 outputs, the attenuation scale of ultrasound and the magnetoacoustic voltage. In this study, an artificial neural network is developed to automate the output of HMM for breast cancer classification. The ANN employs the steepest gradient descent with momentum back propagation algorithm with logsig and purelin transfer function. The best ANN architecture of 3-2-1 (3 network inputs, 2 neurons in the hidden layer, one network output) with learning rate of 0.3, iteration rate of 20000 and momentum constant of 0.3 was successfully developed with accuracy of 90.94% to testing data and 90% to validation data. The result shows the advantages of HMM outputs in providing a combination of bioelectric and acoustic information of tissue for a better breast cancer diagnosis consideration. The system’s high percentage of accuracy shows that the output of HMM is very useful in assisting diagnosis. This additional capability is hoped to improve the existing breast oncology diagnosis.

Keywords—Breast Cancer, Hybrid Imaging, Artificial Neural Network

I. INTRODUCTION

A. Breast Cancer

Breast cancer is one of the most common cancers and the leading cause of cancer death among women worldwide [1, 4]. Breast cancer incidence is increasing over the years with more than 1 million new cases reported each year [3]. In addition to that, an average of 373000 women died globally every year in conjunction to the disease [2, 4]. In Malaysia specifically, National Cancer Registry report for the year 2003-2005 states that, the incidence rate of breast cancer in Malaysian population is 47.4 per 100000 populations with Chinese is at the highest rate of 59.9, followed by Indian at 54.2 and Malay at 34.9 per 100000 populations [10].

Breast cancer is a disease of uncontrolled breast cells growth, in which the cells acquire genetic alteration that allows them to multiply and grow outside the context of normal tissue development [36]. The cell metabolism increases to meet the requirements of rapid cell proliferation, autonomous cell growth and to maintain its survival [36]. The most common symptom of breast cancer is the presence of painless and slowly growth lump that may alter the contour or size of the breast [37-38]. It is also characterized by skin changes, inverted nipple and bloodstained nipple discharge [37-38]. The lymphatic nodes under the armpit may be swollen if affected by cancer. In late stage, the growth may ulcerate through the skin and infected [37-38]. Bone pain, tenderness over the liver, headaches, shortness of breath and chronic cough may be an indication of the cancer spreading to other organs in the body [37].

In the cancerous tissue, changes in density occur due to uncontrolled cell multiplication [36, 39-40, 42], excessive accumulation of protein in stroma [5-6, 41] and enhancement of capillary density [43, 44-46]. On the other hand, changes in conductivity also occur due to increase cellular water and electrolyte content as well as altered membrane permeability and blood perfusion to support metabolism requirements [5-6, 38, 41]. With the yearly increasing trend of breast cancer, improvement in diagnosis and treatment method is desirable to increase survival rate.
B. Ultrasonography in Breast Oncology Diagnostics

In the world of medical diagnostic, breast ultrasonography has an established and significant role in diagnostic of breast abnormalities [47]. Ultrasound is superior from mammography for its non-ionizing radiation. This makes ultrasound an imaging of choice to manage symptomatic breast in younger women as well as in pregnant and lactating mother whom the theoretical radiation of mammography is pertinent [11]. Ultrasonography is also a reliable modality for solid and cystic breast anomaly differentiation [12-13, 15, 47-49]. It is also used in imaging augmented and inflamed breast. However, in the current practice, the proportion of patient in whom breast ultrasonography is considered necessary is only 40% [15]. This means that, ultrasonography is not indicated for the rest 60% of patients referred for breast imaging [15]. This practice explains major constraint of ultrasonography in breast imaging that limits its usage for diagnostic of breast symptoms and for screening asymptomatic patients [11-15, 50-52].

The major problem of ultrasonography is its low sensitivity in detecting small and preinvasive breast cancers [11-15] from normal tissues due to the overlapping ultrasonic characteristics in these tissues [17-18, 33]. Breast ultrasound diagnostic relies on several sonographic features that are based on margin, shape and echotecture. Breast cancers are often characterized by poorly defined margins, irregular borders, spiculation, marked hyperechogenicity, shadowing and duct extension [11].

Another limitation of ultrasound is its inability to detect microcalcification, a calcium residue found in the breast tissue as an early indicator of certain type of breast cancer [12]. In ultrasonography, the presence of microcalcification in tissue is often masked with breast tissue heterogeneity and grainy noise due to speckle phenomena [53-54]. The reasons make microcalcification detection with ultrasonography unreliable [12].

In addition to that, previous study [52] reported that the sensitivity of ultrasonography for breast cancer detection is not only complicates by the low sensitivity of the ultrasound itself but also by the dependency of ultrasound result to operator. This means that, a single sonographic image may be interpreted differently by different operators and the result is relative to the operator skills and experience, variations in human perceptions of the images, differences in features used in diagnosis and lack of quantitative measures used for image analysis [55].

Due to the limitation of the existing ultrasound imaging system in breast cancer detection, an ultrasound-based hybrid system called the Hybrid Magnetoacoustic Method (HMM) has been developed in this study. The hybrid system was designed to be capable to access the acoustic and electric properties of tissue to get better diagnostic information.

C. Hybrid Magnetoacoustic Method

Biological tissue is a conductive element due to the presence of random charges that support cell metabolism [57]. Propagation of ultrasound wave inside the breast tissue will cause the charges to move at high velocity due to the back and forth motion of the wave [55-60]. Moving charges in the present of magnetic field will experience Lorentz Force. Lorentz Force separates the positive and negative charges, producing an externally detectable voltage [55-57, 59] that can be collected using a couple of skin electrodes [55-57, 59].

Following the equation of wave propagation presented by Wen et al [63-65], the magnetoacoustic voltage amplitude is:

\[ V = WRcB_0 \int_{\text{Soundpath initial}}^{\text{Soundpath final}} \frac{\partial}{\partial z} \left( \frac{\sigma(z)}{\rho(z)} \right) p(z, t) \, dt \]  

(1) [63-65]

in which, W is the ultrasound beam width, R is the resistance of the detection circuit , B is the magnetic field intensity, \( \sigma \) is the samples conductivity, \( \rho \) is the samples density and \( p \) is the ultrasound pressure.

According to the equation, the amplitude of magnetoacoustic voltage is not directly proportional to the tissue conductivity, but is also weighted by the tissue density [63-65]. This is due to the fact that density is among the tissue parameters that determine the ultrasound attenuation and eventually reduce the velocity of ionic motion inside the tissue. For a heterogeneous tissue like breast cancer, the conductivity weighted density \( \sigma/\rho \), is nonzero at tissue interface, giving rise to magnetoacoustic voltage. Hence, the magnetoacoustic voltage gives information that is relative to conductivity changes across interface [63-65] and this information can be used to access bioelectric properties of the breast tissue.

In addition to the magnetoacoustic voltage, HMM sensed back the ultrasound wave that is initially used to excite tissue ionic motion for tissue acoustic evaluation and gives information with regards to the acoustic attenuation scale of the breast tissue.

Therefore, HMM is designed to be capable to access the conductivity and density of breast tissues. The block diagram of HMM is shown by figure 1.1.
A series of experiment and quantification on the output of HMM to 24 normal and 25 cancerous mice breast tissue show that the combination of acoustic and bioelectric properties is a promising way of breast cancer diagnostic. The result shows that normal mice breast tissues experience the highest attenuation (2.329±1.103 dBmm⁻¹) followed by cancerous tissue (1.76±1.08 dBmm⁻¹) with the difference of 0.569±0.023dB. In addition to that, mean magnetoacoustic voltage results for tissue the normal and cancerous tissue group are 0.42±0.16 µV and 0.8±0.21 µV respectively. Therefore, this expansion of study on the development of breast cancer diagnosis system by using hybrid magnetoacoustic method and artificial neural network will concentrate on developing an automated system for HMM outputs based on the experimental result.

D. Artificial Neural Network

Artificial intelligence has been used very extensively in modeling biomedical application. It has been proposed as reasoning tool to support clinical decision making since the earliest day of computing. An artificial neural network (ANN) is a nonlinear and complex computational mathematical model for information processing with architectures inspired by neuronal organizational biology [19-21]. The underlying reason for using an artificial neural network in preference to other likely methods of solution is its ability to provide a rapid solution. Depending on the type of problem being considered, ANN is a proven method which is a capable of providing fast assessment and accurate result [19-21]. This is because; ANN works in a simulated parallel manner and is not limited by the serial requirements of the normal program such as in expert system and conventional programming [19-22].

The most valuable property of ANN is its ability to learn and to generalize [22]. Generalization refers to the capability of neural network to produce reasonable outputs for input which is not encountered during training [19, 23]. These attributes mark neural network out from other computational methods [22]. Neural network consists of a number of simple and highly connected processors. Like the brain, ANN can recognize pattern, manage data and most significantly, learn [19-20]. Previous study also showed that, ANN with at least one hidden layer of computational unit is capable of approximating any finite function to any degree of accuracy as a universal approximator [24].

Artificial Neural Network

In medicine, ANN is widely used for modelling, data analysis and diagnostic classifications [19-21, 24-25]. The most common ANN model used in clinical medicine is the multilayer perceptron (MLP) [25]. The most widely used connection pattern is three layer back propagation neural network which have been proved to be useful in modelling input–output relationship [69-70, 77] while the most commonly used transfer functions are linear, log sigmoid and tan sigmoid [78].

The most commonly used indicator of ANN performance is Mean Squared Error (MSE). MSE is the average of the squares of the difference between each output and the desired output. Research performed in [19-20, 23, 25-26, 28, 61] used MSE as the measurement of ANN performance. In addition to that, researches conducted in [23, 61, 29-31] were using the accuracy of the tested data as one of the performance indicator of the ANN. By using this method, network is trained using a part of the data and the remainder is assigned as the testing and validation data.

II. METHODOLOGY

The HMM system is capable to produce 2 outputs: The attenuation scale of ultrasound and the magnetoacoustic voltage. In this study, ultrasound attenuation scale and magnetoacoustic voltage data from 24 normal and 25 cancerous mice breast tissue samples were used. The breast tissues were harvested from a set of tumor bearing mice FVB/N-Tg MMTV PyVT 634 Mul that carries invasive adenocarcinoma and its control strain FVB/N that carries normal breast tissue.

A. Ultrasound Attenuation Scale

The ultrasound attenuation scale is the degree of weakening of ultrasound amplitude and intensity as it propagates through a medium. In HMM, the ultrasound
attenuation scale is determined following the insertion loss method. During the ultrasound measurement, vegetable oil was used as medium for ultrasound propagation and the tissue was immersed into the oil that is located between the ultrasound transmitter and receiver. To measure the attenuation scale of ultrasound in tissue, the power spectral density (PSD) of ultrasound wave propagating through the oil medium was first calculated, followed by the PSD of ultrasound propagating through the oil and tissue. The final attenuation was determined by calculating the difference of PSD between the oil medium and oil medium with tissue at 9.8MHz. The total number of attenuation scale used in the development of breast cancer diagnosis system is presented by Table 2.1.

<table>
<thead>
<tr>
<th>No</th>
<th>Attenuation scale</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cancerous tissue</td>
<td>106</td>
</tr>
<tr>
<td>2</td>
<td>Normal tissue</td>
<td>106</td>
</tr>
</tbody>
</table>

Table 2.1: Number of ultrasound attenuation scale at 9.8MHz

B. Magnetoacoustic Voltage Data

The magnetoacoustic voltage data is the voltage that results from Lorentz Force interaction to conductive charges in tissue. Its amplitude is influenced by tissue conductivity and tissue attenuation level. In HMM, magnetoacoustic voltage measurement was made by touching the tissue surface from the x direction by using the skin electrodes. The total number of magnetoacoustic voltage signals that were recorded in this experiment is presented by Table 2.2.

<table>
<thead>
<tr>
<th>No</th>
<th>Magnetoacoustic voltage signal</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Magnetoacoustic voltage of Normal Breast Tissue (1&amp;2)</td>
<td>212</td>
</tr>
<tr>
<td>2</td>
<td>Magnetoacoustic voltage of Cancerous Breast Tissue (1&amp;2)</td>
<td>212</td>
</tr>
</tbody>
</table>

Table 2.2: Total number of magnetoacoustic voltage signal recorded by HMM.

C. Data Massaging

Data massaging involves restructuring the range of neural network input and target values between zero to one. Massaging is done due to the fact that neural network works best when all its input and output vary within the range of 0-1 [19-20].

D. Data Sampling

Data comprises of 25 cancerous tissues and 24 normal tissues were arranged randomly into the training, testing and validation data shown by Table 2.3. For the ultrasound data, measurement was repeated 5 times for every sample. On the other hand, for magnetoacoustic voltage data, measurement was repeated 10 times.

<table>
<thead>
<tr>
<th></th>
<th>Ultrasound data</th>
<th>Magnetoacoustic voltage data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Side 1</td>
<td>Side 2</td>
</tr>
<tr>
<td>Training</td>
<td>N</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>68</td>
<td>69</td>
</tr>
<tr>
<td>Testing</td>
<td>N</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>27</td>
</tr>
<tr>
<td>Validation</td>
<td>N</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>106</td>
<td>106</td>
</tr>
</tbody>
</table>

*N: Normal, C: Cancerous

Table 2.3: ANN data sampling

E. Development of Artificial Neural Network

The diagnosis of breast cancer in this study was performed by employing a Multilayer Feed Forward Neural Network (MFNN) with 2 inputs. It was trained by using the steepest descent with momentum back propagation algorithm with logsig and purelin transfer function in Matlab environment. The back propagation algorithm is the most commonly used algorithm in medical computational application as were experimented by [19-20, 25, 29].

The measurement of ANN performance was observed by using the Mean Squared Error (MSE) and total classification accuracy of the network to the testing data. Training is best when the ANN is capable to achieve lowest MSE value.

In addition to that, each ANN configuration was tested by using the testing group data to obtain the overall classification accuracy as were experimented in [19-20, 29-30]. By using this method, network was trained using a part of the data and the remainder was assigned as testing and validation data.

The overall flow chart for the development of MFNN in this study is shown by figure 2.1.
III. RESULT

A. Results for the determination of optimized number of neuron in hidden layer.

During the hidden neuron optimization step, the number of neuron in the hidden layer was varied from 1 to 12 while the other parameters including learning rate, iteration rate and momentum constant were fixed to a predetermined value of 0.3, 20000 and 0.2 respectively. From figure 3.1, it can be observed that the lowest MSE value that was achieved during the neuron optimization was given by network with 12 hidden neurons at 0.098. However, its classification accuracy is only 89.09%. On the other hand, the highest classification accuracy of 90.94% was given by 2 hidden neurons with MSE of 0.111. An increase of 0.013 in MSE has given 2% increments in classifications accuracy. Hence, hidden layer size of 2 was chosen for optimum number of neuron in the hidden layer for the ANN with slight compensation in higher MSE value.

B. Result for the determination of optimized learning rate

The ANN learning rate was varied from 0.1 to 0.9 during the learning rate optimization step. Training epochs and momentum constant were kept at their predetermined value of 20000 and 0.2 respectively. The number of neuron in the hidden layer was set to its optimum value of 2 neurons. Figure 3.2 shows that the lowest MSE value was achieved with learning rate of 0.1 to 0.8. In addition to that, learning rate of 0.3 to 0.8 gives the highest classification accuracy of 90.94%. In this case, 6 optimum values of learning rate were obtained with the same MSE value and classification accuracy to testing data. Hence, the best learning rate was chosen by testing the robustness of the ANN to validation data. The validation result shows that learning rate of 0.3 gives 90% accuracy compared to learning rate of 0.4 to 0.8 that gives 85% validation accuracy. Hence, learning rate of 0.3 was chosen as the best learning rate values.
C. Result for the determination of optimized iteration rate

The iteration rate was varied from 5000 to 50000 with a constant increment of 5000 during the learning rate optimization process. Other parameters were set at their optimum value except momentum constant, which was at its predetermined value of 0.2. Figure 3.3 shows that iteration rate which is too short produces ANN with high MSE and low classification accuracy. This indicates that the iteration rate is insufficient to allow the network to converge. It also shows that iteration rate of 20000, 40000 and 50000 give the lowest MSE value of 0.111 and the highest classification accuracy of 90.94%. Hence, iteration rate of 20000 was chosen since this architecture produces lowest MSE and highest classification accuracy at the shortest time interval. In addition to that, this architecture was tested to validation data and it gives the highest validation classification accuracy of 90%.

![Figure 3.3: Iteration rate vs Mean Squared Error (MSE) and the total performance accuracy (%)](image)

The final architecture of ANN for classifications of breast cancer in this study is 3-2-1 (3 network inputs, 2 neurons in the hidden layer, one network output) with learning rate of 0.3, iteration rate of 20000 and momentum constant of 0.3.

C. Result on the determination of optimized momentum constant

The final step for the determination of optimum ANN was to find the best momentum constant for the network. The momentum constant was varied from 0.1 to 0.9 while other parameters were kept at their optimum values. Figure 3.4 indicates that momentum constant of 0.2, 0.3 and 0.5 produces network with lowest MSE value. Among that, momentum constant of value 0.3 and 0.5 gives the highest classification accuracy on testing data. Hence, validation data was used to test the robustness of the system and the result indicates that momentum constant of 0.3 gives highest validation classification accuracy of 90%.

![Figure 3.4: Momentum constant vs Mean Squared Error (MSE) and total performance accuracy (%).](image)

The final classification performance of the optimum ANN for testing and validation data is shown in Table 3.1. The result indicates that the ANN is capable to achieve 90.94% and 90% classification accuracy for testing and validation data. This result shows the advantages of HMM output in providing additional bioelectric parameter of tissue for breast cancer diagnosis consideration. The system’s high percentage of accuracy shows that the output of HMM is very useful in assisting diagnosis. This additional capability is hoped to improve the existing breast oncology diagnosis.

<table>
<thead>
<tr>
<th>Data</th>
<th>Testing Data</th>
<th>Validation Data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Cancer</td>
</tr>
<tr>
<td>Actual Data</td>
<td>28</td>
<td>27</td>
</tr>
<tr>
<td>ANN Result</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>% Accuracy</td>
<td>89.29</td>
<td>92.59</td>
</tr>
<tr>
<td>% Total Accuracy</td>
<td>90.94</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Classification Result of the Neural Network
IV. CONCLUSION

ANN with architecture of 3-2-1 (3 network inputs, 2 neurons in the hidden layer, one network output) with learning rate of 0.3, iteration rate of 20000 and momentum constant of 0.3 was successfully developed with accuracy of 90.94% to testing data and 90% to validation data. The result shows the advantages of HMM outputs in providing a combination of bioelectric and acoustic information of tissue for breast cancer diagnosis consideration.

References:

[37] Breast cancer, Malaysia Oncology Society, malaysianoncology.org, 21st May 2011.
Maheza I. M. Salim was born in Selangor, Malaysia, on 26th August 1983. She received an honour degree in Biomedical Engineering (with Distinction) from The University of Malaya, Malaysia in 2006. Currently, she is pursuing her Ph.D degree in Medical Imaging in Universiti Teknologi Malaysia with specialization in hybrid imaging modality for breast cancer detection.

Ismail Ariffin was born in Penang, Malaysia, on the 13th September 1962. He received a Bachelor of Science degree in Electronic and Electrical Engineering from University of Miami, USA in 1987 and a master degree in Electrical Engineering (Specialization in Electronic) from Universiti Teknologi Malaysia in 1991. Currently he is a senior lecturer at the Faculty of Electrical Engineering, Universiti Teknologi Malaysia. He joined the Faculty of Electrical Engineering at Universiti Teknologi Malaysia since 1983. He is currently a Malaysian Standard (SIRIM) technical committee on Anaesthetic/Respiratory and Electromechanical Devices. His current research interests include medical electronic, medical imaging, digital signal processing and image processing.

Abd Hamid Ahmad was born in Melaka, Malaysia, on the 14th July 1960. He received an honours degree in Electrical and Electronic Engineering in1986 from University of Miami and a postgraduate honours degree in Electronic in1991, from Universiti Teknologi Malaysia. He joined the Electronics Department, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, as Assistant Lecturer in 1987, teaching Electronics, Digital Electronics, Digital System, Microprocessor and Electronic System. His current research interests include digital system design, electronics system design and computer assisted learning.

Bustanur Rosidi was born in Nganjuk, Indonesia on 9th February 1966. He received his Diplom-Biochemiker (degree and MSc in biochemistry) at University of Hannover, Germany in 1994. Between 1995 and 1997 he worked as researcher on a project sponsored by DFG at Department of Molecular Neuro Biochemistry, Ruhr-University of Bochum, Germany. He pursued his PhD at Department of Molecular Genetics University of Hannover between 1997 and 2001, where he graduated in July 2001. His PhD work focused on immune biochemical and molecular biological characterization of antigens originating from plants. In 2002 he joined the group of G. Iliakis at Institute of Medical Radiation Biology, University Medical School, University of Duisburg-Essen, in Essen Germany, as a post doctoral researcher working on the biochemistry of DNA repair, especially on the mechanism of non-homologous end joining (NHEJ). In November 2007 he joined the University Technology Malaysia (UTM) in Johor Bahru, Malaysia, where he is currently working as lecturer at the Institute of Bioproduct Development (IBD).

Prof. Eko Supriyanto is a biomedical engineer. He is head of Diagnostics Research Group, Universiti Teknologi Malaysia. He obtained his PhD in medical electronics from University of Federal Armed Forces, Hamburg, Germany. His research interest is engineering application in medicine. He has more than 130 international publications in the area of medical electronics, medical image processing and medical computing. He has 22 national and international innovation awards and more than 10 patents of biomedical products.