# Artificial neural network to aid differentiation of malignant and benign breast masses by ultrasound imaging

Song, Jae, Venkatesh, Santosh, Conant, Emily, Cary, Ted, Arger, Peter, et al.

Jae H. Song, Santosh S. Venkatesh, Emily F. Conant M.D., Ted W. Cary, Peter H. Arger M.D., Chandra M. Sehgal, "Artificial neural network to aid differentiation of malignant and benign breast masses by ultrasound imaging," Proc. SPIE 5750, Medical Imaging 2005: Ultrasonic Imaging and Signal Processing, (12 April 2005); doi: 10.1117/12.595295

SPIE.

Event: Medical Imaging, 2005, San Diego, California, United States

# Artificial Neural Network to aid differentiation of malignant and benign breast masses by ultrasound imaging

Jae H. Song[a], Santosh S. Venkatesh[a], Emily. F. Conant[b], Ted W. Cary[b], Peter H. Arger[b], Chandra M. Sehgal[b]

[a]Department of Electrical Engineering, [b]Department of Radiology
University of Pennsylvania
Philadelphia, Pennsylvania 19104

## ABSTRACT

The goal of this study is to evaluate an Artificial Neural Network (ANN) for differentiating benign and malignant breast masses on ultrasound scans. The ANN was designed with three layers (input, hidden and output layer), where a sigmoidal (hyperbolic tangent) response function is used as an activation function at each unit. Data from 54 patients with biopsy-proven malignant (N=20) and benign (N=34) masses were used to evaluate the diagnostic performance of the ANN. Of the seven quantitative features extracted from ultrasound images, only four showed statistically significant difference between the two categories. These features were margin sharpness, margin echogenicity, angular continuity, and age of patients. The diagnostic performance was evaluated by round-robin substitution to negate bias due to small sample size. All the input features were standardized to zero-mean and unit-variance to prevent non-uniform learning, which can generate unwanted error. The outputs of the network were analyzed by Receiver Operating Characteristics (ROC). The resulting area under the ROC curve $A_z$ was 0.856 ± 0.058 with 95% confidence limit from 0.734 to 0.936, providing 76.5% specificity at 95% sensitivity. The performance of the ANN was comparable to the performance by logistic regression analysis reported by our group earlier. These results suggest that an ANN when combined with sonography can effectively classify malignant and benign breast lesions.

**Keywords –** Ultrasound imaging, Artificial Neural Network, computer-aided diagnosis, breast cancer.

## 1. INTRODUCTION

Breast cancer is one of the most common cancers in women.[1] More than 175, 000 American women are diagnosed with breast cancer each year. In the last several years, various imaging modalities including X-ray mammography, MRI, PET, optical and ultrasound imaging have been used to detect and diagnose breast cancer.

Sonography has been traditionally used for the detection and diagnosis of cysts. Improvements in image quality from new ultrasound techniques such as compound and harmonic imaging have expanded their role, so that now sonography is commonly used as a complement to X-ray mammography for imaging solid masses. Several attempts have been made in the past to characterize breast tissue by complex signal and image processing schemes. Despite these efforts, interpretation of breast sonograms continues to be done almost exclusively by qualitative assessment of lesion features within an image. A variety of image features based on echogenicity, echotexture, and lesion shape are used, including: hypoechogenicity relative to the surrounding fibroglandular tissue; inhomogeneous echo architecture; an echogenic rim representing desmoplasia; irregular tumor-tissue border, described by the terms spiculation, microlobulation and angular margin; and duct extensions and branching patterns.[2,3] Taken together, these qualitative descriptors have been shown to improve the specificity of clinical and mammographic findings.[4,5] The overall goal of this study is to quantify various features that an observer sees in an image so as to use them with classical and artificial neural networks (ANNs) for tissue classification. This article presents the results of supplying an ANN with several margin characteristics of solid breast nodules.



Medical Imaging 2005: Ultrasonic Imaging and Signal Processing,
edited by William F. Walker, Stanislav Y. Emelianov, Proc. of SPIE Vol. 5750
(SPIE, Bellingham, WA, 2005) · 1605-7422/05/$15 · doi: 10.1117/12.595295

## 2. MATERIALS AND METHODS

Anti-radial and radial ultrasound images of 54 patients who had excisional or core biopsy of suspicious breast masses were selected from a library of images. The studies were approved by the institutional review committee.

Seven mass features were analyzed in the study: 1) Margin Sharpness, 2) Margin Echogenicity, 3) Angular Continuity, 4) Tissue Attenuation, 5) Mass Attenuation, 6) Excess Attenuation, and 7) Patient Age.[6] Except for patient age, the features quantitatively measure shape and margin characteristics of the mass. All of the features were then input into the ANN for analysis. Both the computer programs for feature extraction and the ANN were developed using the IDL language (Research System Inc., Denver, Colorado).

### Feature analysis

Arithmetic mean and standard deviation of each feature were derived for the malignant and benign groups. The two-tailed Student t-test of unequal variance was used to determine the statistical significance of the difference between the two groups. Four out of seven extracted features showed a significant difference between the benign and malignant groups. Only these four features were used as inputs for the ANN analyses. These features included patients' age (Age) and three ultrasound image features: margin sharpness (M-Sharp), margin echogenicity (M-Echo), and angular variance in margin (AVM). M-Sharp represents the diffuseness of the border of a mass; M-Echo represents the difference of brightness (measured in difference in mean grayscale level) between the inside and outside of a mass at the margin; and AVM quantifies the inhomogeneity in margin brightness with angle.

### Artificial Neural Network (ANN)

A multilayer perceptron model of the ANN using a sigmoid function was constructed. The network consisted of an input layer, a hidden layer, and an output layer. The input layer contained four units (neurons) corresponding to four input features; the hidden layer contained ten units transforming the input features from four-dimensional to ten-dimension space. Finally, the output layer had only one neuron, representing two possible diagnostic states: malignant or benign.

The ANN analysis was a two-step process involving training on known examples followed by testing on unknown samples. The training procedure itself consisted of two processes involving feed-forwarding the input data followed by back propagation of error by adjusting weights to minimize error on each training epoch.

### Feed-forwarding

All the features in the training set were standardized with zero-mean and unit-variance. This standardization was necessary to prevent non-uniform learning, in which the weights associated with some features converge faster than others. After standardization, a randomly chosen sample was put into the network for feed-forwarding. For each sample, four features were used as inputs, and these inputs were processed from the hidden layer to the output layer. After each sample was passed through the network, the output value was calculated as

$$z = g(\mathrm{x}) = f\left(\sum_{j=1}^{10} w_{1j} f\left(\sum_{i=1}^{4} w_{ji} x_i\right)\right), \tag{1}$$

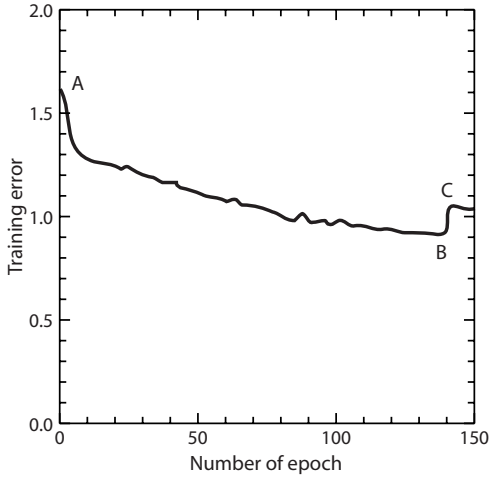
where $f(\cdot)$ is an n-unipolar sigmoid function with a ramp slope of 0.3. The nonlinear nature of the sigmoid function confines the output values between 0 and 1. In Equation 1, $z$ is the output value, $x_i$ are input features, and $w_{ji}$ are weighting values used in the hidden and output layer. After calculating the output value $z$ of the network, the output value was compared to the target value (assigned 1 for malignant and 0 for benign) and the difference was used to determine the training error. This process was repeated for every sample in each training epoch.

**Learning – error back propagation**

After feed-forwarding the entire training set, weights of the network were adjusted to meet the minima of the mean square error criterion function, defined as $J(\mathrm{w}) = \frac{1}{2}\|\mathrm{t} - \mathrm{z}\|^2$, where t and z are vectors of target and output values, respectively. To minimize this criterion function, the weight adjustment value was calculated by $\Delta \mathrm{w} = -\eta \frac{\partial J}{\partial \mathrm{w}}$, where $\eta$ is the user-input learning rate, which was set to 0.1 based on the results of general training.

As a first step, the general behavior of the network was evaluated using all data to confirm that the network could be trained. This also yielded the network parameters, such as learning rate, and the slope of the sigmoid function of the optimum training curve. The optimum curve was defined as a curve that has smooth monotonic decreasing slope and saturates to the minimum training error, like the curve in Figure 1. The parameters obtained from general training were applied to individual cycles of round-robin training and testing.

Fig 1: Training of the ANN. Training curve ABC shows a monotonic decrease in training error. Beyond B, the network becomes over-trained.




**ROC Analysis**

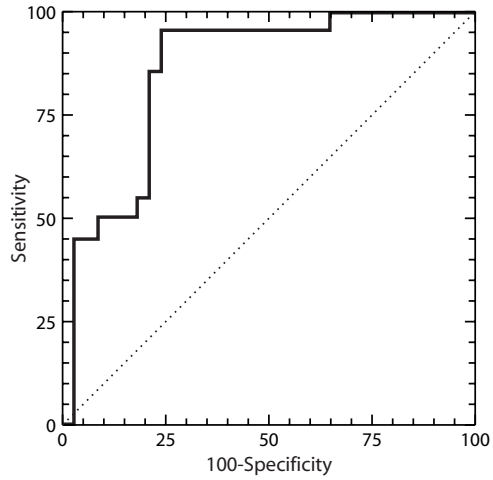
The ANN test outputs were evaluated for diagnostic performance by ROC analysis. The area under the curve, $A_z$, was used as a measure of the diagnostic performance. Specificity was measured at fixed 95% sensitivity.

## 3. RESULTS

In both the general training and the testing, a batch algorithm – adjusting weights after feed-forwarding the entire training set – was used as the learning algorithm. The training curve from the general training is shown in Figure 1. 150 training epochs were used. The training error was minimized and saturated at around 140 epochs. The overtraining occurs after 140 epochs, which is marked by an increase in training error (Fig. 1).

The ROC curve for the ANN is shown in Figure 2. 40 training iterations were used for each test sample. The number of training iterations is different from the general training to minimize the processing time as well as to avoid overtraining problems. The area under the ROC curve, $A_z$, was 0.856 ± 0.058 with 95% confidence limit from 0.734 to 0.936, providing 76.5% specificity at 95% sensitivity.

Fig 2: Graph showing the ROC curve for margin analysis from the Artificial Neural Net.



## 4. DISCUSSION

There is growing interest in using computer-aided analysis of ultrasound images to differentiate between malignant and benign masses.[7-14] Clinically, a variety of image features is used to make this distinction.[2,3] From the many features used, tumor-tissue margin plays an important role in the diagnosis. In this study, we measured margin sharpness, margin echogenicity, and angular continuity of the tumor-tissue interface. Each of these features has been shown to be statistically different for the malignant and benign masses.[6] The mean age of the patients for the two groups was also statistically different. When the three ultrasound image parameters and age of the patients were used with a round-robin ANN for tumor classification, a significant performance was achieved. The outputs of the network analyzed by Receiver Operating Characteristics (ROC) resulted in an area under the ROC curve $A_z$ of 0.856±0.058 with 95% confidence limit of 0.734 to 0.936. At a fixed sensitivity of 95%, 76.5% specificity was observed. These results are promising, especially if one considers that, from the many potential image features, only margin characteristics were selected in this study. Adding new clinical features to the ANN should further improve the performance level. Furthermore, in this study the breast lesions were manually traced, a tedious approach that introduces observer variability into the measurements. An automated method to delineate tumor lesions with high levels of confidence and detail could potentially overcome this limitation.[14-16]

In summary, the results of this study suggest that an ANN when combined with quantitative tumor-tissue margin features can effectively differentiate between malignant and benign breast lesions. Further development of the proposed approach could result in a promising technique for computer-aided diagnosis of breast masses using ultrasound imaging.

## 5. ACKNOWLEDGMENTS

This research was supported in part by NIH grants CA-85424 and CA-87526.

## 6. REFERENCES

1. Wingo PA, Tong T, Bolden S, Cancer Statistics, A Cancer Journal for Clinicians, 41(1): 8-30. 1995.

2. Madjar H. The Practice of Breast Ultrasound: Techniques, Findings Differential Diagnosis. Thieme Medial Publishers, Inc. New York. 2000.

3. Leucht D, Madzjar H. Teaching Atlas of Breast Ultrasound. Thieme Medical Publisher Inc., New York. 1996.

4. Stavros AT, Sonography of solid breast Nodules benign or malignant. The Leading Edge in Ultrasound, 69-84, Thomas Jefferson University, Atlantic City, New Jersey, 1998.

5. Stavros AT. Thickman D. Rapp CL. Dennis MA. Parker SH. Sisney GA. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. *Radiology. 196(1):123-34, 1995.*

6. Sehgal CM, Cary TW, Kangas SA, Weinstein SP, Schultz SM, Arger PH, Conant EF. Computer-based margin analysis of breast ultrasound for differentiating malignant and benign masses, J Ultrasound in Medicine, 23:2001-1209, 2004

7. Goldberg V, Manduca A, Ewert DL, Gisvold JJ, Greenleaf JF. "Improvement in specificity of ultrasonography for diagnosis of breast tumors by means of artificial intelligence," *Medical Physics. 19(6):1475-81, 1992.*

8. Garra BS, Krasner BH, Horii SC, Ascher S, Mun SK, Zeman RK, "Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis", *Ultrasonic Imaging. 15(4):267-85, 1993.*

9. Dumane VA, Shankar PM, Piccoli CW, Reid JM, Forsberg F, Goldberg BB, "Computer aided classification of masses in ultrasonic mammography", *Medical Physics. 29(9):1968-73, 2002.*

10. Shankar PM, Dumane VA, Piccoli CW, Reid JM, Forsberg F, Goldberg BB, "Classification of breast masses in ultrasonic B-mode images using a compounding technique in the Nakagami distribution domain", *Ultrasound in Medicine & Biology", 28(10):1295-300, 2002.*

11. Drukker K, Giger ML, Horsch K, Kupinski MA, Vyborny CJ, Mendelson EB, "Computerized lesion detection on breast ultrasound", *Medical Physics. 29(7):1438-46, 2002.*

12. Chen DR. Chang RF, Kuo WJ, Chen MC, Huang YL, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks", *Ultrasound in Medicine & Biology. 28(10):1301-10, 2002.*

13. Sehgal CM, Arger PH, Rowling SE, Conant EF, Reynolds C, Patton JA, "Quantitative vascularity of breast masses by Doppler imaging: regional variations and diagnostic implications" *Journal of Ultrasound in Medicine, 19(7):427-40; 441-2, 2000.*

14. Chen CM, Chou YH, Han KC, Hung GS, Tiu CM, Chiou HJ, Chiou SY, "Breast lesions on sonograms: computer-aided diagnosis with nearly setting-independent features and artificial neural networks", *Radiology, 226(2):504-14, 2003.*

15. Cary TW, Conant EF, Arger PH, Sehgal CM, “Diffuse boundary extraction of breast masses on ultrasound by leak plugging”, (submitted for publication).

16. Madabhushi A, Metaxas DN, "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions", *IEEE Trans. Med. Imaging, 22(2): 155-169, 2003.*