Computer-Aided Diagnosis for Surgical Office-Based Breast Ultrasound

Ruey-Feng Chang, PhD; Wen-Jia Kuo, MS; Dar-Ren Chen, MD; Yu-Len Huang, PhD; Jau-Hong Lee, MD; Yi-Hong Chou, MD

Hypothesis: The computer-aided diagnostic system is an intelligent system with great potential for categorizing solid breast nodules. It can be used conveniently for surgical office-based digital ultrasonography (US) of the breast.

Design: Retrospective, nonrandomized study.

Setting: University teaching hospital.

Patients: We retrospectively reviewed 243 medical records of digital US images of the breast of pathologically proved, benign breast tumors from 161 patients (ie, 136 fibroadenomas and 25 fibrocystic nodules), and carcinomas from 82 patients (ie, 73 invasive duct carcinomas, 5 invasive lobular carcinomas, and 4 intraductal carcinomas). The digital US images were consecutively recorded from January 1, 1997, to December 31, 1998.

Intervention: The physician selected the region of interest on the digital US image. Then a learning vector quantization model with 24 autocorrelation texture features is used to classify the tumor as benign or malignant. In the experiment, 153 cases were arbitrarily selected to be the training set of the learning vector quantization model and 90 cases were selected to evaluate the performance. One experienced radiologist who was completely blind to these cases was asked to classify these tumors in the test set.

Main Outcome Measure: Contribution of breast US to diagnosis.

Results: The performance comparison results illustrated the following: accuracy, 90%; sensitivity, 96.67%; specificity, 86.67%; positive predictive value, 78.38%; and negative predictive value, 98.11% for the computer-aided diagnostic (CAD) system and accuracy, 86.67%; sensitivity, 86.67%; specificity, 86.67%; positive predictive value, 76.47%; and negative predictive value, 92.86% for the radiologist.

Conclusion: The proposed CAD system provides an immediate second opinion. An accurate preoperative diagnosis can be routinely established for surgical office-based digital US of the breast. The diagnostic rate was even better than the results of an experienced radiologist. The high negative predictive rate by the CAD system can avert benign biopsies. It can be easily implemented on existing commercial diagnostic digital US machines. For most available diagnostic digital US machines, all that would be required for the CAD system is only a personal computer loaded with CAD software.


Breast ultrasound (US) has become an increasingly integral part of the evaluation, diagnosis, and treatment of breast disease. It is the most useful adjunctive technique to mammography. Also, it plays an important role in differentiating cystic from solid masses and in guiding interventional procedures. With US, the levels of diagnostic confidence and accuracy depend on the quality of the examination. Good-quality equipment must be used to produce high-quality image resolution. The skilled operator must properly mark the region of interest (ROI) to achieve the correct diagnosis and differential diagnosis. The rapid development of US made it seem advisable to reconsider the clinical value of breast US, especially using the high-resolution, real-time US, and computer-aided diagnostic (CAD) system. High-resolution probes, computer-enhanced imaging, and portable machinery have led to the widespread adoption of real-time US by breast surgeons. Surgeons have a much greater clinical correlation than radiologists by performing the digital US studies. Breast surgeons should not be excluded in the multidisciplinary care of a patient with breast disease. On the contrary, it must include the surgeon, radiologist, pathologist, and patient. Ultrasonographic examination is painless, requires...
PATIENTS AND METHODS

The US examination was performed using a handheld 7.5-MHz lineal array transducer; no acoustic standoff pad was used. All solid masses measured from 0.8 to 3.7 cm were sampled percutaneously by fine needle aspiration or open biopsy and then correlated with pathological features. We retrospectively reviewed 243 medical records of digital US images of the breast of pathologically proved, benign breast tumors from 161 patients (ie, 136 fibroadenomas and 25 fibrocystic nodules), and carcinomas from 82 patients (ie, 73 invasive duct carcinomas, 5 invasive lobular carcinomas, and 4 intraductal carcinomas). The images were consecutively recorded from January 1, 1997, to December 31, 1998. One breast surgeon (D.-R.C.) captured all digitized US images and sampled all aspiration specimens. The patients’ ages ranged from 17 to 64 years (average age, 42 years).

The image of the ROI is manually selected with a region that extended beyond the lesion margins by 1 to 2 mm in all directions and then saved as a file for later analysis by the proposed LVQ system. A real-time digitized US monochrome image and an ROI subimage of a tumor is shown in the Figure. The texture correlation between neighboring pixels within an ROI subimage is used to classify the tumor. We adopted the normalized autocorrelation coefficients as the texture features of a tumor. The 2-dimensional normalized autocorrelation coefficient between pixel \((i, j)\) and pixel \((i+\Delta m, j+\Delta n)\) in an image with size \(M\times N\) is defined as follows:

\[
\gamma(\Delta m, \Delta n) = \frac{A(\Delta m, \Delta n)}{A(0, 0)},
\]

where

\[
(2) \quad A(\Delta m, \Delta n) = \frac{1}{(M-\Delta m)(N-\Delta n)} \sum_{x=-\Delta m}^{\Delta m} \sum_{y=-\Delta n}^{\Delta n} |f(x, y) - \bar{f}| |f(x+\Delta m, y+\Delta n) - \bar{f}|.
\]

and \(\bar{f}\) is the mean value of \(f(x, y)\). To remove the influence of brightness on US images, the mean value is removed from each pixel’s value.

In this study, the LVQ1 learning algorithm has been selected to train the proposed LVQ model. In this work, we use the LVQ Program Package (LVQ_PAK, Version 3.1; prepared by the LVQ Programming Team of the Helsinki University of Technology Laboratory of Computer and Information Science, Helsinki, Finland, April 7, 1995). The detailed descriptions of the learning algorithms could be found in the program document of LVQ_PAK. The 2-dimensional normalized autocorrelation matrix is used as the input of the LVQ model. The dimension of the matrix can be fixed for any size of image. In this work, a US ROI image produces a \(5\times5\) autocorrelation matrix as the texture features. Notices that the value of \(\gamma(0, 0)\) is always 1 for a normalized autocorrelation matrix. Hence, except for the element \(\gamma(0, 0)\), other autocorrelation coefficients are formed as a 24-dimensional image feature vector. To diminish the occurrence of dead neuron in the LVQ model, we remove the dead neurons according to the training set after a number of training iterations. Meanwhile, the ambiguous codeword will be split with the members of the training set in the cluster to create 2 new code words. The steps will be performed until the amount of code words reaches an acceptable percentage of the training set.

Recently, technical advances in US have expanded the potential usefulness of this modality for the evaluation of breast lesions. Scientists were less familiar with neural networks. In early 1989, the Joint Conference on Neural Networks, conducted in San Diego, Calif, helped to make this field grow considerably. A neural network comprises computer programs that learn to make diagnostic predictions. The construction is simple. The neural networks learn by example and they deal with ambiguous data. They can provide an educated guess, something conventional computer programs could not do well, or at all. Previous articles have suggested that neural networks can assist physicians with diagnosis. They have been applied to the prediction of breast cancer and ovarian malignancy. They are potentially useful for differential diagnosis of interstitial lung disease. However, the aforementioned studies applied the database of clinical findings. Texture features are helpful to classify masses and normal tissue on mammograms and the most useful features were those derived from co-occurrence matrices of the images. In this study, the physician located the ROI of digital US images and used the autocorrelation features of a tumor to make the diagnosis by using the learning vector quantization (LVQ) neural networks.

The improvement of classification performance can be achieved by connecting with supervised learning.
The arbitrarily selected 153 digital US images (ie, 101 benign breast tumors and 52 carcinomas) are used to train the LVQ. The test set contains 30 carcinomas and 60 benign breast tumors. To compare the performance of the proposed CAD system with the radiologist, an experienced radiologist (Y.-H.C.) (familiar with breast US examinations) who was completely blind to these test cases was asked to classify these tumors on a 3-point scale (where 1 indicates benign; 2, indeterminate; and 3, malignant). The category of “not benign” is the number including malignant and indeterminate cases of the sonographic classification according to Stavros et al.15 This represents the total number of lesions requiring biopsy according to their US classification. Table 1 lists the classification of breast nodules by the proposed LVQ and an experienced radiologist.

For the LVQ learning algorithms, the maximal number of iterations is limited to 10000 and the initial codebook sizes 10, 20, 30, 40, and 50 are selected. We find that the effective and similar performance is achieved by using codebook size in the ranges 20 to 30. Using an initial codebook size of 20 and limit the amount after splitting codewords to be 10% of the training set, the LVQ correctly identifies 29 of 30 malignant tumors and 52 of 60 benign tumors. Table 2 lists the summary of performance between the LVQ diagnostic system and the radiologist (Y.-H.C.). The results illustrated the following: accuracy, 90%; sensitivity, 96.67%; specificity, 86.67%; positive predictive value, 78.38%; and negative predictive value, 98.11% for the CAD system and accuracy, 86.67%; sensitivity, 86.67%; specificity, 86.67%; positive predictive value, 76.47%; and negative predictive value, 92.86% for the experienced radiologist.

### Table 1. Classification of Breast Nodules by the Radiologist and the Proposed LVQ*

<table>
<thead>
<tr>
<th>Sonographic Classification</th>
<th>Radiologist†</th>
<th>LVQ‡</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Benign</td>
<td>Malignant</td>
</tr>
<tr>
<td>Benign</td>
<td>52</td>
<td>4</td>
</tr>
<tr>
<td>Indeterminate</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Malignant</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>30</td>
</tr>
</tbody>
</table>

*LVQ indicates learning vector quantization; TN, true negative; TP, true positive; FP, false positive; and FN, false negative. All are confirmed histological findings.
†The accuracy of the radiologist was as follows: 52 benign tumors were diagnosed correctly in 60 cases (TN = 52); 4 carcinomas were diagnosed as benign (FN = 4); 26 carcinomas were diagnosed as not benign in 30 cases (TP = 26); and 8 benign tumors were categorized as not benign (malignant + indeterminate) (FP = 8).
‡The accuracy of the LVQ was as follows: 52 benign tumors were diagnosed correctly in 60 cases (TN = 52); 1 carcinoma was diagnosed as benign (FN = 1); 29 carcinomas were diagnosed as not benign in 30 cases (TP = 29); and 8 benign tumors were categorized as not benign (FP = 8).

### Table 2. Summary of Performance Between the Radiologist and the LVQ* Diagnostic System

<table>
<thead>
<tr>
<th>Item</th>
<th>Radiologist</th>
<th>LVQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy, %†</td>
<td>86.67</td>
<td>90.00</td>
</tr>
<tr>
<td>Sensitivity, %‡</td>
<td>86.67</td>
<td>96.67</td>
</tr>
<tr>
<td>Specificity, %§</td>
<td>86.67</td>
<td>86.67</td>
</tr>
<tr>
<td>Positive predictive value %¶</td>
<td>76.47</td>
<td>78.38</td>
</tr>
<tr>
<td>Negative predictive value %¶</td>
<td>92.86</td>
<td>98.11</td>
</tr>
</tbody>
</table>

*LVQ indicates learning vector quantization; TP, true positive; TN, true negative; FP, false positive; and FN, false negative.
†Accuracy = (TP + TN)/(TP + TN + FP + FN).
‡Sensitivity = TP/(TP + FN).
§Specificity = TN/(TN + FP).
¶Positive predictive value = TP/(TP + FP).
*Negative predictive value = TN/(TN + FN).

Interpretive criteria established for breast lesions have been accepted as appropriate descriptors for the interpretation of breast abnormality. The experienced radiologists, Stavors et al,15 and Skaane and Engedal,16 reported that the sensitivity of breast US for malignancy was 98.4% and 99.55%, the specificity was 67.8% and 29%, the positive predictive value was 38% and 66%, and the negative predictive value was 99.5% and 98%, respectively; the overall accuracy as reported by Stavros et al was 72.9%. If they are well trained, surgeons will be expected to have the same accuracy rate of interpretation as experienced radiologists. This is not a hard task for surgeons because the surgeon is the primary care provider for the patient. Surgeons learn the performance of US easily because they do the cytology or biopsy of patients by themselves so that they can correlate the information of both image and abnormality. However, because some features overlap between benign and malignant criteria, surgeons may sometimes have difficulty in interpreting findings with confidence. The proposed CAD programs are an intelligent system with extreme potential for categorizing solid breast nodules with a high accuracy rate, which was compared with the results of experienced radiologists. Computer-aided diagnosis of radiological images has become a rapidly expanding field of research. Methods of image texture analysis are undergoing great development and use within the field of medical imaging. The combination of image texture analysis and automated decision making of the LVQ model provides an immediate second opinion and an accurate preoperative diagnosis can be routinely established for surgical office-based digital US of breast. It can be easily implemented on existing commercial diagnostic digital US machines. For most available diagnostic digital US machines, all that would be required for the CAD system is only a personal computer loaded with CAD software.
REFERENCES


ARCHIVES OF INTERNAL MEDICINE

Blood Levels of Homocysteine and Increased Risks of Cardiovascular Disease: Causal or Casual?
William G. Christen, ScD; Umed A. Ajani, MBBS; Robert J. Glynn, ScD; Charles H. Hennekens, MD

Background: Accumulating data from epidemiological studies suggest that individuals with elevated blood levels of homocysteine have increased risks of cardiovascular disease. We reviewed the currently available evidence of an association between homocysteine and cardiovascular disease and examined whether the strength of the evidence varies according to study design.

Methods: We used a computerized MEDLINE literature search, 1966 through September 1998, to identify all epidemiological studies that examined the relationship of homocysteine level with risks of coronary heart or cerebrovascular disease. Two measures of plasma homocysteine level and its association with risk of cardiovascular disease were extracted: mean homocysteine level in cases and controls, and relative risk of cardiovascular disease for elevated homocysteine level.

Results: A total of 43 studies were reviewed. Most cross-sectional and case-control studies indicated higher mean homocysteine levels (either fasting or after methionine load) and/or a greater frequency of elevated homocysteine level in persons with cardiovascular disease as compared with persons without cardiovascular disease. Results of most prospective studies, however, indicated smaller or no association. The few prospective studies that reported a positive association between homocysteine level and risks of cardiovascular disease included patients with preexisting vascular disease.

Conclusions: In contrast to cross-sectional and case-control studies, results of prospective studies indicated less or no predictive ability for plasma homocysteine in cardiovascular disease. Instead, elevated homocysteine level may be an acute-phase reactant that is predominately a marker of atherogenesis, or a consequence of other factors more closely linked to risks of cardiovascular disease. Randomized trials are necessary to test reliably whether lowering homocysteine levels will decrease risks of cardiovascular disease. (2000;160:422-434)

Reprints: William G. Christen, ScD, Division of Preventive Medicine, Brigham and Women’s Hospital and Harvard Medical School, 900 Commonwealth Ave E, Boston, MA 02215-1204.