*Technology in Cancer Research and Treatment*
*ISSN 1533-0346*
*Volume 10, Number 4, August 2011*


# Cost-Effective and Non-Invasive Automated Benign & Malignant Thyroid Lesion Classification in 3D Contrast-Enhanced Ultrasound Using Combination of Wavelets and Textures: A Class of ThyroScan™ Algorithms

**U. R. Acharya, Ph.D., D.Eng.[1]**
**O. Faust, Ph.D.[1]**
**S. V. Sree, Ph.D.[2,*]**
**F. Molinari, Ph.D.[3]**
**R. Garberoglio, M.D.[4]**
**J. S. Suri, MS, Ph.D., MBA[5]**

www.tcrt.org

[1]Dept. of ECE, Ngee Ann Polytechnic, Singapore 599489
[2]Researcher Room 3, N3-01c-10, School of Mechanical and Aerospace Engineering, College of Engineering, Nanyang Technological University, 50, Nanyang Avenue, Singapore 639798
[3]Biolab, Department of Electronics, Politecnico di Torino, Torino, Italy
[4]Scientific Foundation Mauriziana Onlus, Torino, Italy
[5]Fellow AIMBE, CTO, Global Biomedical Technologies Inc., Roseville, CA, USA & Biomedical Engineering Department, Idaho State University (Aff.), ID, USA

Ultrasound has great potential to aid in the differential diagnosis of malignant and benign thyroid lesions, but interpretative pitfalls exist and the accuracy is still poor. To overcome these difficulties, we developed and analyzed a range of knowledge representation techniques, which are a class of ThyroScan™ algorithms from Global Biomedical Technologies Inc., California, USA, for automatic classification of benign and malignant thyroid lesions. The analysis is based on data obtained from twenty nodules (ten benign and ten malignant) taken from 3D contrast-enhanced ultrasound images. Fine needle aspiration biopsy and histology confirmed malignancy. Discrete Wavelet Transform (DWT) and texture algorithms are used to extract relevant features from the thyroid images. The resulting feature vectors are fed to three different classifiers: K-Nearest Neighbor (K-NN), Probabilistic Neural Network (PNN), and Decision Tree (DeTr). The performance of these classifiers is compared using Receiver Operating Characteristic (ROC) curves. Our results show that combination of DWT and texture features coupled with K-NN resulted in good performance measures with the area of under the ROC curve of 0.987, a classification accuracy of 98.9%, a sensitivity of 98%, and a specificity of 99.8%. Finally, we have proposed a novel integrated index called *Thyroid Malignancy Index (TMI)*, which is made up of texture features, to diagnose benign or malignant nodules using just one index. We hope that this *TMI* will help clinicians in a more objective detection of benign and malignant thyroid lesions.

Key words: Thyroid lesion; Computer Aided Diagnosis; Contrast Enhanced Ultrasound; Texture; Discrete Wavelet Transform.

## *Introduction*

The National Cancer Institute estimated the number of new thyroid cancer cases to be 44,670 and the predicted number of deaths due to this cancer to be 1,690 in 2010 (1). Another database documents a recent rise in the number of cases of thyroid carcinoma, with an estimated increase of 3% per year in the incidence of thyroid cancer (2, 3). Although the incidence of thyroid cancer appears to be increasing,

**Abbreviations:** AUC: Area under the ROC curve; CAD: Computer Aided Diagnosis; CEUS: Contrast-Enhanced Ultrasound; DeTr: Decision Tree; DICOM: Digital Imaging and Communications in Medicine; DUS: Doppler Ultrasound; DWT: Discrete Wavelet Transform; FN: False Negatives; FNA: Fine Needle Aspiration; FNAB: Fine Needle Aspiration Biopsy; FP: False Positives; GLCM: Gray Level Co-occurrence Matrix; HRUS: High Resolution Ultrasound; K-NN: K-Nearest Neighbor; PNN: Probabilistic Neural Network; RAM: Random Access Memory; RI: Resistance Index; ROC: Receiver Operating Characteristic; TMI: Thyroid Malignancy Index; TN: True Negatives; TP: True Positives.

[*]Corresponding author:
S. V. Sree, Ph.D.
E-mail: vinithasree@ntu.edu.sg
jsuri@comcast.net

the number of patients evaluated with a thyroid nodule without carcinoma remains far greater. Thyroid nodules are very common and may occur in more than 50% of adult population with about 7% of thyroid nodules being diagnosed as malignant (4). In general, the incidence of thyroid nodules is on the rise due to the wider use of neck imaging (5).

Such statistics indicate that there is an urgent need for cost-effective thyroid diagnosis support systems. Cost efficiency is important because a large number of tests must be performed in order to detect a relatively small number of cancer cases. In terms of diagnosis technique, Fine Needle Aspiration (FNA) biopsy is considered to be the "gold standard" in the diagnosis of thyroid nodules (6). However, FNA is too labor intensive to be used for large scale screenings and has many pitfalls (7). A more effective detection strategy is to analyze medical images, because in this case the diagnosis process can be automated. Such systems are called as Computer Aided Diagnosis (CAD) systems. Use of thyroid images for the diagnosis is possible, because the image texture indicates the histopathologic components of the thyroid nodules. Now, having established the possibility of using thyroid images for cancer detection, we need to address the issue of cost efficiency. There is a wide range of medical imaging modalities available which could be used for thyroid nodule diagnosis. Ultrasound imaging is, by a large margin, the most cost effective of these modalities, and many studies have demonstrated that ultrasound characteristics of thyroid nodules can predict the risk of malignancy (8, 9).

There is a wide range of different ultrasound imaging methods in existence. One of them is frequency encoded Doppler Ultrasound (DUS) imaging which has been used for the identification of flow in thyroid tumors. However, the role of DUS in the evaluation of thyroid nodules for malignancy has not yet been accurately studied (10). Internal flow without or with minimal peripheral flow on DUS and Resistance Index (RI) $\geq 0.70$ were used to distinguish between malignant and benign thyroid nodules reliably. Nodules with prevailing peripheral vascularisation and minimal or no internal vascularisation, and RI below 0.70 were found to be probably benign. Doppler studies need a quantitative evaluation of the internal nodule flow to avoid subjective interpretations and partial visions caused by the bidimensional nature of the traditional High Resolution Ultrasound (HRUS). The sonographic features like the size and echogenicity of the tumors, the presence of cystic areas or calcifications, and detectable blood flow on color Doppler imaging of Hürthle Cell Neoplasms (HCNs) of the thyroid were studied (11). They concluded that the Hürthle cell neoplasms showed a spectrum of sonographic appearances from predominantly hypoechoic to hyperechoic lesions and from peripheral blood flow with no internal flow to extensively vascularized lesions. They also indicated that the differentiating benign and malignant HCNs was difficult using ultrasound and FNA techniques, and therefore, complete removal of the lesion is the only safe option.

Contrast-Enhanced Ultrasound (CEUS) imaging was introduced to enhance the differential diagnosis of solitary thyroid nodules. The feasibility of CEUS imaging of the thyroid gland and the potential of this method for characterizing solitary thyroid nodules were studied (12). They assessed the baseline echogenicity and the dynamic enhancement pattern of each nodule, in comparison with adjacent thyroid parenchyma. Their results show that CEUS of thyroid gland was a feasible technique. However, overlapping findings seem to limit the potential of this technique in the characterization of thyroid nodules. Recently, enhancement patterns of thyroid nodules on gray-scale contrast-enhanced ultrasound were evaluated for the differential diagnosis (13). Their results show that CEUS enhancement patterns were different in benign and malignant lesions. The application of ultrasonographic contrast agents that lead to an improvement in the differential diagnosis of thyroid nodules was studied (14). In the group of benign lesions, in the patients affected by nodular goiter, an intra-nodular perfusion as opposite to the healthy surrounding parenchyma was observed. Even though the ultrasound contrast agent technique has a limited invasivity and is more expensive than FNA, the preliminary data of this pilot study suggested that this method may be useful to differentiate benign from malignant thyroid nodules.

In this study, we have used Discrete Wavelet Transform (DWT) and texture based feature extraction methods for the differential diagnosis of malignant and benign thyroid lesions. We used these features from CEUS images because this imaging method is more efficient in differentiating benign from malignant lesions. The following section presents the materials and methods used in the development of the CAD system (a class of ThyroScan™ algorithms from Global Biomedical Technologies, Inc., California, USA) for feature extraction, classification, and statistical analysis. In the subsequent *results* section, both classification results and associated statistical analysis are discussed. We also propose an integrated diagnostic index called the *Thyroid Malignancy Index (TMI)* in this section. The literature presented in this introduction section forms the basis for a discussion. The *discussion* section projects the proposed thyroid diagnosis technique into a wider perspective by comparing the results of the proposed system with previously published classification results. We also discuss on the cost-effectiveness of the proposed technique and present its limitations. In the *conclusion* section of this paper, we highlight the accuracy of the proposed method.

### *Materials and Methods*

Figure 1 shows the block diagram of the proposed CAD system (a class of ThyroScan™ algorithms from Global

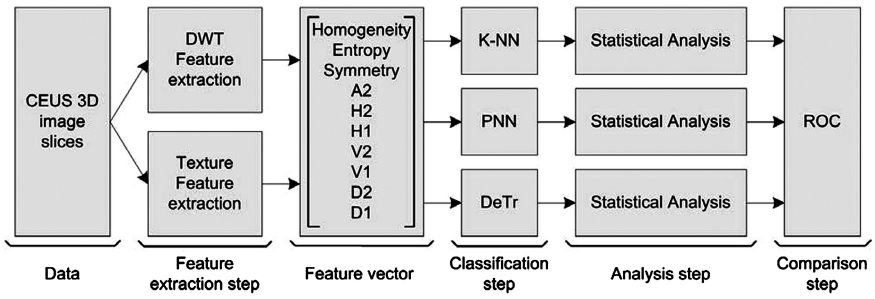


**Figure 1:** Block diagram of the proposed CAD system (a class of ThyroScan™ algorithms).

Biomedical Technologies, Inc., California, USA). In general, computer aided diagnosis systems can be constructed with a feature extraction subsystem and a classification subsystem. In this work, we used both DWT and texture based features. The extracted feature vectors were fed to one of the three classifiers: K-Nearest Neighbor (K-NN), Probabilistic Neural Network (PNN), and Decision Tree (DeTr). The role of each individual component in the block diagram is described in this section.

*Data*

Twenty patients with previously confirmed diagnosis of solitary thyroid nodule were enrolled in this study. Ten subjects were male (age: 53.5 $\pm$ 13.3 years; range: 22 – 71 years) and ten were female (age: 50.1 $\pm$ 10.8 years; range: 25 – 68 years). All the patients signed an informed consent prior to participating in the experiment. The experimental protocol was approved by the ethical committee of the Endocrinology Section of the "Umberto I'" Hospital of Torino (Italy).

All subjects underwent a clinical examination, hormonal profiling, and ultrasound (B-Mode and Color Doppler) examination of the lesion. Then, 2.5 ml of ultrasound contrast agent (Sonovue, Bracco, Italy) was administered intravenously and a 3-D volume containing the lesion was acquired. Due to bulkiness and weight of external mechanical scanning systems and the variability associated with the nodules dimension and its position, we preferred to perform a free-hand scanning. A trained operator with more than 30 years of experience in neck ultrasonography (R.G.) performed all the scans. The high frame rate of the device compared to the slow movement of the probe ensured that there was no gap between adjacent frames. The average frame rate of the device during acquisition was 16 Hz. Images were acquired by a MyLab70 ultrasound scanner (Biosound-Esaote, Genova, Italy) equipped by a LA-522 linear probe working in the range 4-10 MHz. All the images were acquired at 10 MHz. The volumes were transferred in DICOM format to an external workstation (Apple PowerPc, dual 2.5 GHz, 8 G RAM) equipped with processing and reconstruction software.

We originally enrolled 30 non-consecutive patients into the study. Five patients were discarded since they had a nodule of more than 6 cm of diameter, about twice the average size of the nodules of our sample set. Manual scanning can be difficult in the presence of big lesions; therefore, we discarded these five patients. Two patients were discarded due to artifacts during acquisition (motion artifact such as swallowing and cough). All the subjects underwent ultrasound-guided Fine Needle Aspiration Biopsy (FNAB) of the thyroid lesion. In the end, three patients were discarded since FNAB documented an ongoing inflammatory state (thyroiditis). Hence, we were left with a population of 20 subjects, ten of which had benign lesions, and ten had malignant lesions. Five nodules were considered benign (goiter nodules) and classified as THY2 according to cytological criteria (15). Of the remaining fifteen patients classified as THY3, five had benign conditions (follicular neoplasm) and ten had malignant conditions (seven papillary, one follicular and two Hurtle cells carcinoma). We acquired 40 data sets from each of the ten patients diagnosed with malignant nodules. Similarly, we acquired 40 data sets from each patient having benign thyroid nodules. Overall, our study contains 400 benign and 400 malignant data sets. The ten patients who were diagnosed with malignant nodules underwent thyroidectomy. The histo-pathological analysis confirmed the diagnosis of malignant carcinoma for all the ten patients. The results of the FNAB were used as reference for the benign nodules: all were struma nodules. Figures 2 and 3 show the typical malignant (papillary carcinoma) and benign (goiter nodule) thyroid sonographic images, respectively: panel A shows the HRUS images, panel B the Color Doppler, and panel C the CEUS. The white arrows indicate the lesion and the red arrow indicates a portion of healthy thyroid parenchyma.

*Feature Extraction*

Feature extraction is one of the most important steps in ThyroScan™, an automated CAD system, because this step extracts relevant and representative features from measurement data such as images and signals. In this work, DWT and texture features were extracted from CEUS images.

*DWT Feature Extraction*

DWT is a useful and efficient tool for many image processing applications. DWT uses filter banks which are composed from finite impulse response filters (16). These filters are used for decomposing signals into low and high

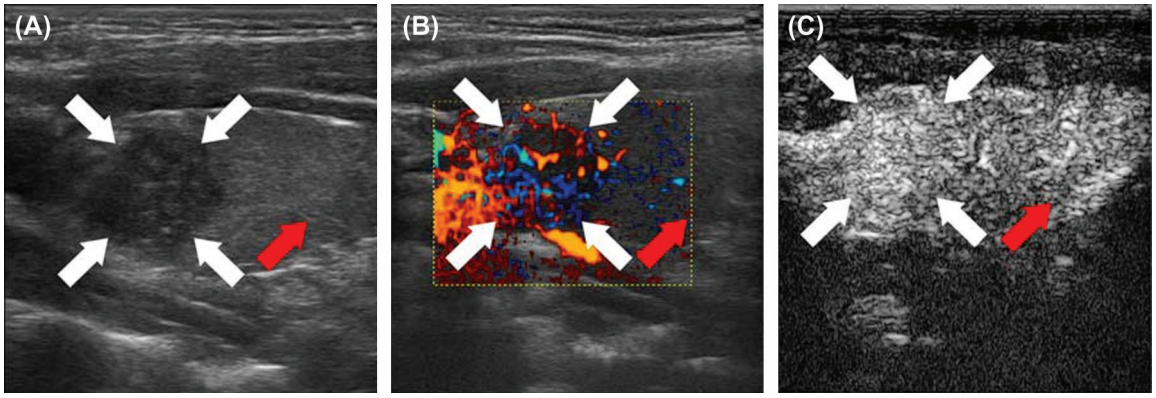


**Figure 2:** Typical appearance of a malignant nodule (papillary carcinoma): **(A)** HRUS; **(B)** Color Doppler; **(C)** CEUS. The white arrows indicate the lesion. The red arrow indicates a portion of healthy thyroid parenchyma.

pass components, called sub-bands. The low pass components contain information (in the form of coefficients) about slow varying signal characteristics, and the high pass components contain information about sudden changes in the signal. When DWT is applied to images, there are four different filtering possibilities:

(i) Low pass filtering is performed on both rows and columns. The resulting LL coefficients contain most of the image's total energy.
(ii) Low pass filtering is performed on the rows, and high pass filtering on the columns. The resulting HL coefficients contain the vertical details of the image.
(iii) High pass filtering is applied to the rows, and low pass filtering to the columns. The resulting LH coefficients contain the horizontal details of the image.
(iv) High pass filtering is conducted on the rows and columns (HH coefficients). The resulting HH coefficients contain the diagonal details of the image, and they are the finest-scale wavelet coefficients.

Decomposition is further performed on the LL sub-band to attain the next coarser scale of wavelet coefficients.

In our work, the CEUS images were first converted to a gray-scale representation and then DWT was applied. Figure 4 shows the complete passband structure for a 2D sub-band transform with three levels of decomposition. The left part of Figure 4 shows the sub-band decomposition of a CEUS image. The right part of Figure 4 represents the extracted features which were calculated as the energies of the individual sub-bands. To be specific, $A_2$, $V_2$, $H_2$, $D_2$, $H_1$, $V_1$ and $D_2$ are the energies of $LL_2$, $HL_2$, $LH_2$, $HH_2$, $LH_1$, $HL_1$ and $HH_1$, respectively. In this work, we have used Daubechies (Db) 8 as the mother wavelet. For this study, we stopped at the $2^{nd}$ level decomposition, because higher level decompositions did not yield significant features. The individual sub-bands are represented as matrixes and these matrixes are combined to form a feature. The method for combining the matrix elements is the same for all sub-band features. All the elements within the individual rows of the matrix are added and the elements of the resulting vector are squared before adding to form a scalar. Finally, this scalar is normalized by dividing it by the number of rows and columns of the original matrix to obtain the energy. $A_2$, $H_2$, $H_1$, $V_2$, $V_1$, $D_2$, and $D_1$ (as shown Figure 4) are equivalent to the notations A2, H2, H1, V2, V1, D2, and D1 (as shown in Figure 1).

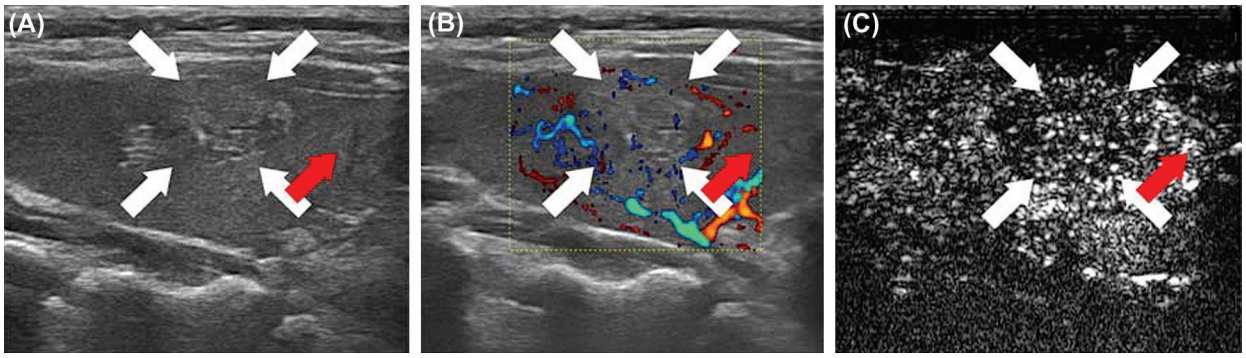


**Figure 3:** Typical appearance of a benign nodule (goiter nodule): **(A)** HRUS; **(B)** Color Doppler; **(C)** CEUS. The white arrows indicate the lesion. The red arrow indicates a portion of healthy thyroid parenchyma.

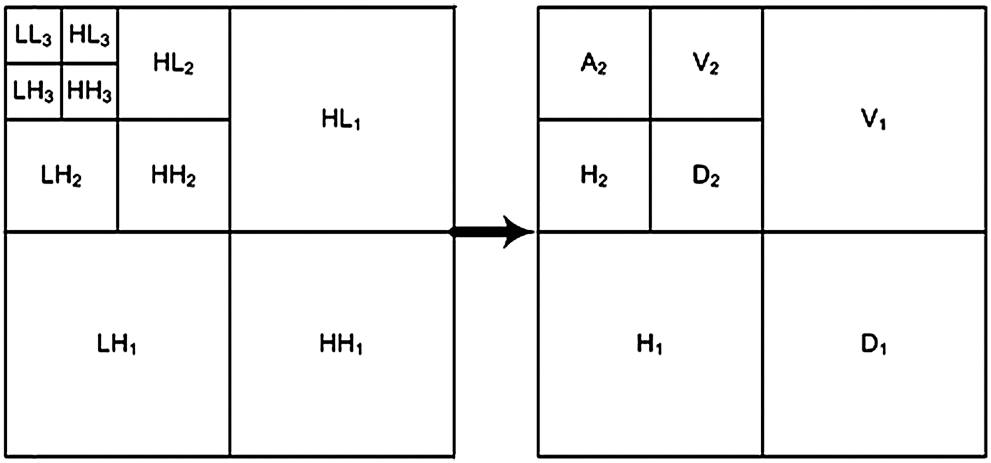


**Figure 4:** Passband structure for a 2D sub-band DWT with three levels.

*Texture Feature Extraction*

Texture features measure smoothness, coarseness, and regularity of pixels which form an image. These measures describe a mutual relationship among intensity values of neighboring pixels repeated over an area larger than the size of the relationship (17). There are two common approaches to texture analysis: statistical analysis and structural analysis. In the statistical approach, scalar measurements of the textures are obtained. This approach characterizes textures as smooth, coarse, or grainy *etc*. These methods are based on both distributions and relationships between intensity values of pixels. Measures include entropy, contrast, and correlation based on the gray level co-occurrence matrix. Structural texture analysis is more complex when compared to the statistical approach (18). It presents detailed symbolic descriptions of the image. Parameters that are extracted using the statistical approach are more suitable for image analysis than those obtained using the structural method (19). In this section, the statistical parameters extracted from the CEUS images are briefly described. The Gray Level Co-occurrence Matrix (GLCM) of an $M \times N$ image $I$ is defined (16) by

$$C_d(i, j) = \left| \{(p, q), (p + \Delta x, q + \Delta y): I(p, q) = i, I(p + \Delta x, q + \Delta y) = j\} \right| \quad [1]$$

where $(p, q), (p + \Delta x, q + \Delta y) \in M \times N$, $d = (\Delta x, \Delta y)$ and $|\cdot|$ denotes the cardinality of a set. The probability of a pixel with a grey level value $i$ having a pixel with a gray level value $j$ at a $(\Delta x, \Delta y)$ distance away in an image is

$$P_d(i,j) = \frac{C_d(i,j)}{\sum_i \sum_j C_d(i,j)} \quad [2]$$

Based on the above mentioned, we obtain the following features:

$$Homogeneity = \sum_i \sum_j \frac{P_d(i,j)}{1+(i-j)^2}, \quad [3]$$

$$Entropy = -\sum_i \sum_j P_d(i,j) \times \ln[P_d(i,j)], \text{ and} \quad [4]$$

$$Symmetry = 1 - \sum_i \sum_j P_d(i,j) - P_d(j,i) \quad [5]$$

*Classification Process*

We employed supervised classifiers in this work. The inputs to these classifiers are the DWT-based and texture-based features that were extracted using the methods described above. In the training phase, the grayscale feature vectors belonging to a set of images in the dataset along with the corresponding class label (benign or malignant) were used in the classifier to allow it to learn the classifier parameters (called the training parameters) that effectively relate the feature vectors and the corresponding class labels. Subsequently, during the testing phase, these training parameters were used on the feature vectors from the remaining images to predict their respective class labels and to calculate the efficiency of the classifier in terms of its classification accuracy. Three classifiers were studied in this work, namely K-Nearest Neighbor (K-NN), Probabilistic Neural Network (PNN), and Decision Tree (DeTr). They are briefly described in this section.

*K-Nearest Neighbor (K-NN)*

K-NN is based on the minimum distance from a query instance to the training samples. The K-nearest neighbors are determined using this method. After gathering these K-nearest neighbors, the majority of them are used for the prediction (20).

*Probabilistic Neural Network (PNN)*

PNN is a specific type of two layer radial basis network which is often used for classification. The first layer of neurons in a PNN has radial basis activation functions. This layer computes the distance vector by evaluating the distances between the input and training vectors. The second layer (competitive layer) sums the contributions of each input classes and produces a vector of probabilities as the output of the input classes. The so called compete transfer function, at the output of the second layer, selects the maximum of these probabilities and assigns a 1 for the selected class and a 0 for all other classes (21).

*Decision Tree (DeTr)*

DeTr classifier generates a tree and a set of rules to represent the model in order to identify different classes from a given data. The rules can be used to recognize the unknown data (22).

*Statistical Analysis*

The student's t-test is a form of regression analysis used to assess whether the two groups have different *means* on some measure. If there is less than 5% chance of getting the observed differences by chance, then a statistically significant difference between the two groups is reported. The lower 'p' values indicate that these groups are clinically significant.

The Receiver Operating Characteristic (ROC) curve is a plot in a two dimensional space. The x-axis is '1 – specificity' and the y-axis is 'sensitivity'. Sensitivity, also known as true positive fraction, refers to the probability that a test result is positive when a disease is present. The Area under the ROC curve (AUC) indicates the classifier performance across the entire range of cut-off points. Conventionally, the area under the ROC curve must fall in the range between 0.5 and 1 (23). An area closer to one indicates that the classifier has a better accuracy. The area under the ROC curve is a good indicator for the classifier performance (24).

## *Results*

The range of features, classification results, and the description of a novel index called the *Thyroid Malignancy Index (TMI)* and its range for the two classes are given in this section. Table I documents the results of statistical analysis of the DWT and texture features. The last column of this table shows the p-value of the features. The fact that all p-values are below 0.0001 indicates that all features are clinically significant. The homogeneity, symmetry and all the DWT features are higher for malignant nodules compared to benign nodules because benign images (Figure 3) have more structure in their sonographic appearance compared to malignant thyroid images (Figure 2). The images with more structure, such as the benign thyroid images, have more variations in the grayscale values compared to the malignant thyroid images, and therefore have higher entropy values.

Ten-fold stratified cross validation method was used to test the classifiers. Using this technique, the whole dataset was split into ten equal parts (roughly), each part containing a similar proportion of images belonging to both classes. Nine parts of the data (training set) were used for classifier development and the built classifier was evaluated using the remaining one part (test set) (*i.e.,* 720 images were used for training and 80 images for testing each time). This procedure was repeated ten times using a different part as the test set in each case. Average of the accuracy, sensitivity, specificity, and AUC obtained during each of the ten folds was calculated to obtain the overall performance measures.

Table II presents the classification results obtained by using the extracted DWT and texture features in the three classifiers. The first column indicates the classifier used. The next four columns present the number of True Negatives (TN), False Negatives (FN), True Positives (TP), and False Positives (FP). The classification accuracy is shown in column 6. Columns 7, 8 and 9 show the sensitivity, specificity, and AUC, respectively. The classification accuracy for all three tested classifiers is well above 96%. It is also evident from the results that the K-NN classifier performs better than DeTr and PNN with a higher accuracy of 98.9%. To be the optimal classifier for a particular dataset, the classifier should not only have a high accuracy, but also a balance between the sensitivity and specificity values. Such a balance is necessary to ensure that the selected classifier classifies samples belonging to both classes with equal error margin and is not biased towards classifying samples belonging to only one class correctly. Therefore, the K-NN classifier was chosen as the optimal classifier for this dataset as it not only presents the highest accuracy of 98.9% but also has equally good sensitivity of 98% and specificity of 99.8%. Figure 5 shows the ROC curves of the three classifiers. It can be clearly seen from the figure below that, the K-NN performs better than the other two classifiers.

*Thyroid Malignancy Index (TMI)*

In this work, we have used Entropy, Homogeneity and Symmetry texture features to develop an integrated index called the *Thyroid Malignancy Index (TMI).* It is difficult to track how these three texture features vary in a patient for making an appropriate diagnosis. Hence, we have formulated an integrated index by combining these features in such a way that the index is distinct for benign and malignant nodules. The *TMI* is defined as follows.

**Table I**
Values of the various DWT and texture features for the benign and malignant thyroid lesions.

| Features | Benign (Mean ± SD) | Malignant (Mean ± SD) | p-value |
|---|---|---|---|
| Homogeneity | 0.285 ± 1.742E-02 | 0.345 ± 4.280E-02 | < 0.0001 |
| Entropy | 3.78 ± 5.402E-02 | 3.52 ± 0.160 | < 0.0001 |
| Symmetry | 0.851 ± 1.355E-02 | 0.860 ± 2.941E-02 | < 0.0001 |
| A2 | 1.635E+06 ± 6.189E+05 | 3.446E+06 ± 2.335E+06 | < 0.0001 |
| H2 | 4.20 ± 2.30 | 5.12 ± 3.06 | < 0.0001 |
| H1 | 0.362 ± 0.237 | 0.474 ± 0.334 | < 0.0001 |
| V2 | 319 ± 124 | 676 ± 665 | < 0.0001 |
| V1 | 66.1 ± 42.6 | 206 ± 214 | < 0.0001 |
| D2 | 1.07 ± 0.626 | 2.20 ± 1.20 | < 0.0001 |
| D1 | 2.907E-02 ± 6.303E-03 | 3.362E-02 ± 1.744E-02 | < 0.0001 |

$$TMI = \frac{Entropy}{Homogeneity \times Symmetry} \quad [6]$$

Such an integrated index would help in a faster and more objective detection of benign and malignant thyroid nodules. It is also easier and cost-effective to incorporate this index into a software and use it for diagnosis. Table III shows the *TMI* values (mean ± standard deviation) for the two classes. It can be seen from the table that they are distinctly different from each other without any overlap. Figure 6 shows the box plot of the mean value of *TMI* indices highlighting the separation between the two classes clearly.

### *Discussion*

The recent advances in ultrasound techniques have paved the way for many investigators to propose new imaging algorithms to diagnose the malignant thyroid lesions. Ultrasound methods for thyroid cancer diagnosis are cost-effective, and these methods perform as good as other thyroid cancer diagnosis methods. In this section, we present the comparison of the results obtained using our technique and other techniques in the literature which also aim to diagnose malignant thyroid nodules.

HRUS imaging revealed that there are specific sonographic features that are common findings in malignancy (*i.e.*, microcalcifications, marked hypoechogenicity, irregular margins and the absence of a hypoechoic halo around the nodule) (8). These features alone, however, cannot be used to perform differential diagnosis. In fact, D'Souza *et al*. (25) measured the HRUS sensitivity and specificity of 80.8% and 81.6%, respectively.

Color and Power Doppler are useful for assessing macro vessels, but they cannot evidence micro vessels, which are typical of malignancy (26). The most important Doppler studies on thyroid lesions were qualitative or, at most, semi-quantitative (10, 27). Also, many studies concluded that Doppler imaging could not be used to exclude malignancy because of poor specificity (28). Lyshchik *et al*. (27) have analyzed the accuracy of quantitative analysis of tumor vascularity on power Doppler sonograms in differentiating malignant and benign solid thyroid nodules using tumor histologic evaluation as the reference standard. Univariate analysis showed that there a statistically significant increase in the levels of intranodular vascularization with an increase in tumor size. With the quantification method of vascular indices they achieved an accuracy of 84.5%. Frates *et al*. (28) based their study on thyroid nodules also on the analysis of color Doppler images. They conclude that solid hypervascular thyroid nodes carry a risk of nearly 42% of being malignant. Furthermore, they state that color characteristics for in Doppler images cannot be used to exclude malignancy.

Recent studies showed that CEUS imaging is a very promising technique for the differential diagnosis of thyroid lesions.

**Table II**
TN, FN, TP, FP, Accuracy, Sensitivity, Specificity, and AUC values for the various classifiers.

| Classifiers | TN | FN | TP | FP | Accuracy (%) | Sensitivity (%) | Specificity (%) | AUC |
|---|---|---|---|---|---|---|---|---|
| K-NN | 40 | 1 | 39 | 0 | 98.9 | 98 | 99.8 | 0.987 |
| PNN | 40 | 2 | 38 | 0 | 97.8 | 95.8 | 99.8 | 0.975 |
| DeTr | 40 | 2 | 38 | 0 | 96.9 | 94 | 99.8 | 0.975 |

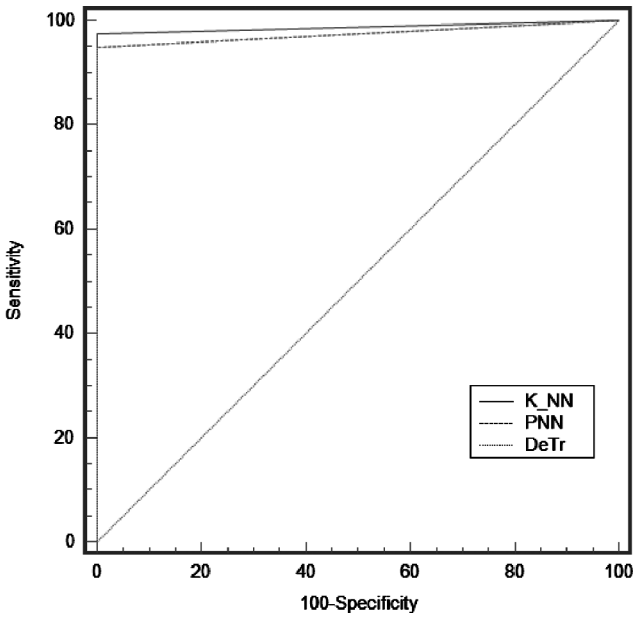


**Figure 5:** ROC curves for the various classifiers of the CAD system (from a class of ThyroScan™ based systems).

In 2010, Zhang *et al*. (13) showed that malignant and benign lesions have different CEUS time-intensity curves. This perfusion study evidenced that ring enhancement was predictive of benign lesions, whereas heterogeneous enhancement was helpful for detecting malignant lesions. This study, however, was based on the computation of perfusion curves by 2-D CEUS imaging. In our study, we used 3-D CEUS imaging in order to utilize the nodule volume. Moreover, in the case of CEUS, the contrast agent enhances the parenchyma representation, given the abundant vasculature of the thyroid gland. Therefore, by using CEUS, it should be possible to gain a better representation of the thyroid vascular pattern, overcoming the limitations given by the low spatial resolution of the Color Doppler imaging.

Finley *et al*. (29) classified benign and malignant thyroid nodules using molecular profiling. In their study, they carried out cluster analysis using 62 samples from two classes (benign and malignant). The results of their study show sensitivity and specificity of 91.7% and 96.2%, respectively. Cerutti *et al*. (30) proposed a pre-operative diagnostic method to distinguish benign and malignant thyroid carcinoma based on gene expression. A total thyroidectomy was the treatment of choice, and a negative result was confirmed on permanent pathology in 20 cases. The immunohistochemistry correctly

**Table III**
Range of *TMI* for benign and malignant cases.

| Features | Benign | Malignant | p-value |
|---|---|---|---|
| *TMI* | 15.6 ± 1.13 | 12.1 ± 1.79 | < 0.0001 |

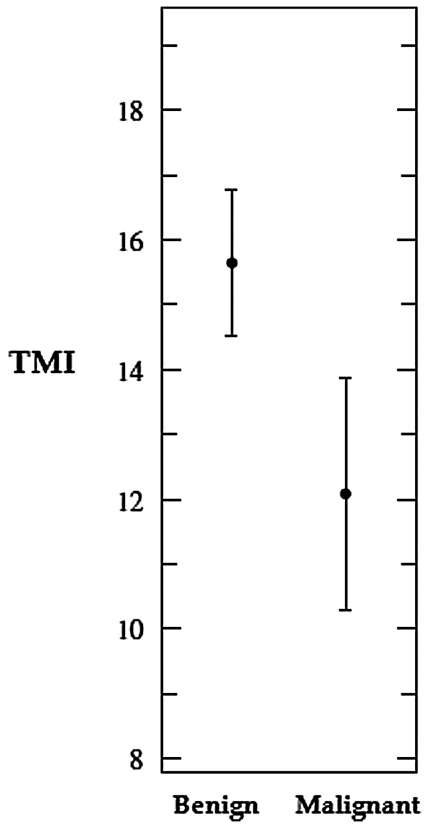


**Figure 6:** Box plot of the *TMI*.

classified 29 of 32 fine-needle aspirations (90.6%) and 23 of 27 follicular thyroid adenomas (85%). The authors of this study were not satisfied with both sensitivity and specificity, and therefore, they proposed further work to increase both measures.

Patton *et al*. (31) differentiated between malignant and benign solitary thyroid nodules by fluorescent scanning. They demonstrated a sensitivity of 93.8% in the identification of cancer. However, the accuracy in the distinction between benign and malignant tissues was only 77.0%. Regardless of the cost, the classification accuracy was low for state of the art CAD systems. The B-mode sonographic images of inflamed and healthy tissues were differentiated automatically using texture features (32). A classification success rate of 100% was achieved with as few as one optimal feature among the 129 texture characteristics tested. The stability of the results with respect to sonograph setting, thyroid gland segmentation and scanning direction was tested. In this work, authors have studied the normal and inflamed ultrasound images.

Based on the above studies, we felt the necessity for a better technique that can improve the classification efficiency and that is also more economical. In our previous study published in 2010 (33), we applied an advanced image processing and

modeling technique to the CEUS 3-D volumes of benign and malignant lesions, and quantified the parameters of the intranodal vascularity. In this paper, which presents the final step of our study pipeline, we developed a computer aided classification framework for the differential diagnosis of thyroid lesions using ultrasound images from 3-D CEUS which can effectively evidence microvessels and intranodular vascularity. We believe that our strategy could be a step forward in the quantitative and user-independent classification of thyroid lesions. Our proposed method is simple and does not involve intensive computation. We used the extracted features in classifiers and concluded that a combination of DWT and texture parameters coupled with a simple K-NN classifier can be used for automated classification with an accuracy, sensitivity and specificity of more than 98%. We believe that this high value of accuracy is due to the choice of the features used in the classifiers. DWT features extract the information from the images in both the time and frequency domains, and hence, these features capture the subtle variations in the characteristics of both the benign and malignant images. In the case of texture features, the homogeneity feature measures the similarity between two pixels that are ($\Delta x$, $\Delta y$) apart. Denseness and degree of disorder in an image are measured by energy and entropy features. In general, the entropy feature will have a maximum value when all elements of the co-occurrence matrix are the same. The symmetry projections indicate prominent directions within the texture of CEUS images, and therefore, symmetry is an important discriminative feature of these images. In addition to this, we have proposed a novel integrated index called *Thyroid Malignancy Index (TMI)* that can be used to identify benign and malignant conditions with high accuracy. This index is distinct for each of the two classes, and therefore, can help in faster, easier, more cost-effective, and more objective detection of benign and malignant lesions.

Finally, we would like to remark that the computer aided diagnosis of malignancy is critical to reduce the number of unnecessary thyroid excisions. Mihai *et al*. (15) showed that the most problematic thyroid lesion is the one classified as THY3 by FNA biopsy and about one fourths of those with THY3 classification had a thyroid carcinoma. THY3 nodules are to be considered suspect even though not malignant. THY3 nodules often show HRUS features that correlate to malignancy and, therefore, are sent to surgical thyroidectomy. This means that about three fourths of the THY3 patients still undergo a surgical excision that could be avoided. Since the cost associated with the CEUS examination followed by the use of the proposed CAD technique is relatively less than the cost of a surgical intervention, this CEUS based classification scheme can be considered as cost-effective. The technique is also minimally invasive compared to surgical thyroidectomy. Moreover, the proposed technique can be programmed into software which can be written in a CD and shipped to hospitals or be downloaded from the internet at no extra cost."

This study has some limitations. The first limitation is given by the relatively low sample size we used. As we discussed in the Materials and Methods section, we had to discard 30% of the patients (10 out of 30) due to inflammation of the thyroid gland (three patients), to the excessive size of the lesion (fie patients), or to artifacts during the acquisition stage (two patients swallowed or coughed). Since only about 7% of the thyroid nodules are malignant (34), building a large database of malignant nodules requires a big effort. We are currently working to enlarge our database. The second limitation is given by the fact that the HRUS and CEUS images were acquired by manual scanning. Mechanical ultrasound linear probes optimized for both HRUS and CEUS are still not easily available. We preferred to use manual instead of mechanically driven scanning because (1) the mechanical device was bulky and not well tolerated by some patients and (2) there are non-negligible costs associated to the motor controlled systems and maintenance.

### *Conclusion*

There is a need for the cost-efficient biomedical diagnostic support systems. In this work, we have investigated the performance of the proposed CEUS based thyroid cancer CAD system using texture and DWT parameters. The extracted DWT and texture features were fed as input to the three different classifiers to compare their performances. Our results show that the combination of DWT and texture features coupled with K-NN classifier presented a classification accuracy of 98.9%, sensitivity of 98% and 99.8% specificity.

In order to make the differentiation faster and more objective, we have gone one step further and formulated a non-dimensional integrated index (given by Eqn. 6) that is composed of texture features. Based on the information presented in Table III and Figure 6, it is evident that this integrated *TMI* Index can be employed for the diagnosis of benign and malignant nodules effectively. The advantage of this Integrated Index is the fact that, in order to make a diagnosis, the physician needs to only look at the value of just one integrated index instead of checking the range of each individual feature. Hence, this *TMI* can be used as an adjunct tool for the clinicians to cross check their diagnosis.

### *Conflict of Interest Statement*

Dr. Jasjit S. Suri has a financial interest in the design of this CAD system (ThyroScan™), but this could not influence the writing or publication of this manuscript. None of the other authors have any financial or personal conflict of interest that

could inappropriately influence the writing or publication of this manuscript.

***References***


1. NCI (National Cancer Institute) on thyroid cancer. Information available at http://www.cancer.gov/cancertopics/types/thyroid (last accessed August 2010).
2. ACS (American Cancer Society). What are the key statistics about thyroid cancer? Information available at http://www.cancer.org/Cancer/ThyroidCancer/DetailedGuide/thyroid-cancer-key-statistics (last accessed August 2010).
3. US CSWG (U.S. Cancer Statistics Working Group). United States Cancer Statistics: 2000 Incidence. Atlanta, GA: Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute; 2003.
4. Lansford, C. D., Teknos, T. N. Evaluation of the Thyroid Nodule. *Cancer Cont 13*, 89-98 (2006).
5. Polyzos, S. A., Kita, M., Avramidis, A. Thyroid nodules – stepwise diagnosis and management. *Hormones (Athens) 6*, 101-119 (2007).
6. Baloch, Z. W., Fleisher, S., LiVolsi, V. A., Gupta, P. K. Diagnosis of "follicular neoplasm": a gray zone in thyroid fine-needle aspiration cytology. *Diagn Cytopathol 26*, 41-44 (2002).
7. Caraway, N. P., Sneige, N., Samaan, N. A. Diagnostic pitfalls in thyroid fine-needle aspiration: a review of 394 cases. *Diagn Cytopathol 9*, 345-350 (1993).
8. Bastin, S., Bolland, M. J., Croxson, M. S., Role of ultrasound in the assessment of nodular thyroid disease. *J Med Imaging Radiat Oncol 53*, 177-187 (2009).
9. Chen, S. J., Yu, S. N., Tzeng, J. E., Chen, Y. T., Chang, K. Y., Cheng, K. S., Hsiao, F. T., Wei, C. K. Characterization of the major histopathological components of thyroid nodules using sonographic textural features for clinical diagnosis and management. *Ultrasound Med Biol 35*, 201-208 (2009).
10. Ivanac, G., Brkljacic, B., Ivanac, K., Huzjan, R., Skreb, F., Cikara, I. Vascularisation of benign and malignant thyroid nodules: CD US evaluation. *Ultraschall Med 28*, 502-506 (2007).
11. Maizlin, Z. V., Wiseman, S. M., Vora, P., Kirby, J. M., Mason, A. C., Filipenko, D., Brown, J. A. Hurthle cell neoplasms of the thyroid: sonographic appearance and histologic characteristics. *J Ultrasound Med 27*, 751-757 (2008).
12. Bartolotta, T. V., Midiri, M., Galia, M., Runza, G., Attard, M., Savoia, G., Lagalla, R., Cardinale, A. E. Qualitative and quantitative evaluation of solitary thyroid nodules with contrast-enhanced ultrasound: initial results. *Eur Radiol 16*, 2234-2241 (2006).
13. Zhang, B., Jiang, Y. X., Liu, J. B., Yang, M., Dai, Q., Zhu, Q. L., Gao, P. Utility of contrast-enhanced ultrasound for evaluation of thyroid nodules. *Thyroid 20*, 51-57 (2010).
14. Appetecchia, M., Bacaro, D., Brigida, R., Milardi, D., Bianchi, A., Solivetti, F. Second generation ultrasonographic contrast agents in the diagnosis of neoplastic thyroid nodules. *J Exp Clin Cancer Res 25*, 325-330 (2006).
15. Mihai, R., Parker, A. J., Roskell, D., Sadler, G. P. One in four patients with follicular thyroid cytology (THY3) has a thyroid carcinoma. *Thyroid 19*, 33-37 (2009).
16. Tan, J. H., Ng, E. Y. K., Acharya, U. R. Study of normal ocular thermogram using textural parameters. *Infrared Phys Techn 53*,120-126 (2009).
17. Mirmehdi, M., Xie, X., Suri, J. S. Hand Book of Texture Analysis. Imperial College Press, UK (2008).
18. Gonzalez, R. C., Woods, R. E. *Digital Image Processing*. Prentice Hall, New Jersey (2001).
19. Castellano, G., Bonilha, L., Li, L. M., Cendes, F. Texture analysis of medical images. *Clin Radiol 59*, 1061-1069 (2004).
20. Dasarathy, B. V. *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques*. Los Alamitos: IEEE Computer Society Press (1990).
21. Specht, D. Probabilistic neural networks. *Neural Networks* 3, 109-118 (1990).
22. Breiman, L., Friedman, J. H., Stone, C. J., Olshen, R. A. *Classification and regression trees*. Chapman & Hall (1984).
23. DeLeo, J. Receiver Operating Characteristic Laboratory (ROCLAB): Software for developing decision strategies that account for uncertainty management in artificial neural network decision-making. In: *Proceedings of Second International Symposium on Uncertainty Modeling and Analysis,* 141-144 (1993).
24. Downey, T. J., Meyer, D. J., Price, R. K., Spitznagel, E. L. Using the receiver operating characteristic to assess the performance of neural classifiers. *Neural Networks 5*, 3642-3646 (1999).
25. D'Souza, M. M., Marwaha, R. K., Sharma, R., Jaimini, A., Thomas, S., Singh, D., Jain, M., Bhalla, P. J., Tripathi, M., Tiwari, A., Mishra, A., Mondal, A., Tripathi, R. P. Prospective evaluation of solitary thyroid nodule on 18F-FDG PET/CT and high-resolution ultrasonography. *Ann Nucl Med*. (in press).
26. Argalia, G., De Bernardis, S., Mariani, D., Abbattista, T., Taccaliti, A., Ricciardelli, L., Faragona, S., Gusella, P. M., Giuseppetti, G. M. Ultrasonographic contrast agent: evaluation of time-intensity curves in the characterisation of solitary thyroid nodules. *Radiol Med 103*, 407-413 (2002).
27. Lyshchik, A., Moses, R., Barnes, S. L., Higashi, T., Asato, R., Miga, M. I., Gore, J. C., Fleischer, A. C. Quantitative analysis of tumor vascularity in benign and malignant solid thyroid nodules. *J Ultrasound Med 26*, 837-846 (2007).
28. Frates, M. C., Benson, C. B., Doubilet, P. M., Cibas, E. S., Marqusee, E. Can color Doppler sonography aid in the prediction of malignancy of thyroid nodules? *J Ultrasound Med 22*,127-131 (2003); quiz 132-134.
29. Finley, D. J., Zhu, B., Barden, C. B., Fahey, T. J. 3rd Discrimination of benign and malignant thyroid nodules by molecular profiling. *Ann Surg 240*, 425-436 (2004).
30. Cerutti, J. M., Delcelo, R., Amadei, M. J., Nakabashi, C., Maciel, R. M. B., Peterson, B., Shoemaker, J., Riggins, G. J. A preoperative diagnostic test that distinguishes benign from malignant thyroid carcinoma based on gene expression. *J Clin Invest 113*, 1234-1242 (2004).
31. Patton, J. A., Hollifield, J. W., Brill, A. B., Lee, G. S., Patton, D. D. Differentiation between Malignant and Benign Solitary Thyroid Nodules by Fluorescent Scanning. *The J Nucl Med 17*, 17-21 (1976).
32. Smutek, D., Sára, R., Sucharda, P., Tjahjadi, T., Svec, M. Image texture analysis of sonograms in chronicinflammations of thyroid gland. *Ultrasound Med Biol 29*, 1531-1543 (2003).
33. Molinari, F., Mantovani, A., Deandrea, M., Limone, P., Garberoglio, R., Suri, J. S. Characterization of single thyroid nodules by contrast-enhanced 3-D ultrasound, *Ultrasound Med Biol 36*, 1616-1625 (2010).
34. Hoang, J. K., Lee, W. K., Lee, M., Johnson, D., Farrell, S. US Features of thyroid malignancy: pearls and pitfalls, *Radiographics 27*, 847-860; discussion 861-865 (2007).


*Received: August 11, 2010; Revised: April 22, 2011; Accepted: May 13, 2011*