34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Ovarian Tumor Characterization and Classification: A class of GyneScan™ Systems

U Rajendra Acharya, Vinitha Sree S*, Luca Saba, Filippo Molinari, Stefano Guerriero, Jasjit S Suri,

Fellow AIMBE

***Abstract* — In this work, we have developed an adjunct Computer Aided Diagnostic (CAD) technique that uses 3D acquired ultrasound images of the ovary and data mining algorithms to accurately characterize and classify benign and malignant ovarian tumors. In this technique, we extracted image-texture based and Higher Order Spectra (HOS) based features from the images. The significant features were then selected and used to train and test the Decision Tree (DT) classifier. The proposed technique was validated using 1000 benign and 1000 malignant images, obtained from 10 patients with benign and 10 with malignant disease, respectively. On evaluating the classifier with 10-fold stratified cross validation, we observed that the DT classifier presented a high accuracy of 95.1%, sensitivity of 92.5% and specificity of 97.7%. Thus, the four significant features could adequately quantify the subtle changes and nonlinearities in the pixel intensities. The preliminary results presented in this paper indicate that the proposed technique can be reliably used as an adjunct tool for ovarian tumor classification since the system is accurate, completely automated, cost-effective, and can be easily written as a software application for use in any computer.**

***Index Terms*—ovarian tumor; texture features; higher order spectra; characterization; classification; computer aided diagnosis**

I. Introduction

Ultrasonography and the determination of the levels of a tumor marker called Cancer-Antigen 125 (CA125) are the most commonly used techniques for detecting ovarian cancer. In the case of ultrasonography, the accuracy and reproducibility of the visual interpretations are most often dependent on the skill of the observer. The CA125 marker has been found to be elevated only in 50% of stage 1 cancers.[1] Furthermore, CA125 can also be raised in other malignancies such as uterine and pancreatic, and sometimes in certain benign conditions.[2] In this work, we have proposed an adjunct diagnostic technique that uses image mining techniques to classify ovarian tumors in ultrasound images, and therefore, to give a valuable second opinion to doctors in order to decide further diagnostic protocol for the patient. Such Computer Aided Diagnostic (CAD) techniques can prove to be excellent adjunct techniques, especially for mass screening, because of their speed, non-invasiveness, easy usability, cost-effectiveness, and reliability.

There are very few studies in the application of CAD for ovarian cancer detection. Most of these studies use features based on (a) blood test results,[3] (b) Mass Spectrometry (MS) data,[4-8] and (c) ultrasound images.[9-13] Such MS based classification studies are affected by the curse of dimensionality[14] as they have to process a high dimensional feature set obtained from a small sample size. Moreover, the MS equipment is expensive and not available in most countries. Therefore, in our work, we have proposed the use of images acquired using the commonly available and low-cost ultrasound modality. 3D ultrasonography approach allows for objective and quantitative documentation of the morphological characteristics of benign and malignant tumors.[15] Studies have shown that the selective use of 3D ultrasonography and power Doppler ultrasound can improve the diagnostic accuracy of ovarian tumors.[16] Therefore, we have used 3D transvaginal ultrasonography for image acquisition in this work.

II. Materials and Methods

A. Methodology

Fig. 1 depicts the block diagram of the proposed real-time image mining CAD technique (a class of GyneScan™ systems). It consists of an on-line classification system which predicts the class label (benign or malignant) of a test image based on the transformation of the on-line grayscale feature vector by the training parameters determined by an off-line learning system. The off-line classification system produces the training parameters using the combination of grayscale off-line training features and the respective off-line ground truth training class labels (0/1 for benign/malignant). The grayscale features for on-line or off-line training are based on image texture and Higher Order Spectra (HOS). Significant features among the extracted ones are selected using the *t*-test. We evaluated the Decision Tree (DT) classifier. The above CAD system was developed using an image database, in the training set images were used to develop the DT classifier. The built classifier was evaluated using the test set. For evaluation, we used a *k-fold*

Manuscript received January 30, 2012.

U. Rajendra Acharya is with the Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore

Vinitha Sree is with Global Biomedical Technologies Inc., CA, USA (email: vinitha.sree@gmail.com)

Luca Saba is with the Department of Radiology, Azienda Ospedaliero Universitaria di Cagliari, Cagliari, Italy

F. Molinari is with the Dept. Electronics and Telecommunications, Politecnico di Torino, Torino, Italy.

Stefano Guerriero is with Department of Obstetrics and Gynecology, University of Cagliari, Ospedale San Giovanni di Dio, Via Ospedale 46, 09124 Cagliari, Italy

Jasjit S. Suri, Fellow AIMBE, is a CTO with Global Biomedical Technologies, CA, USA and is also affiliated with Biomedical Engineering Department, Idaho State University, ID, USA (jsuri@comcast.net).

cross validation protocol. The predicted class labels of the test images and the corresponding ground truth labels (0/1) are compared to determine the performance measures of the system such as sensitivity, specificity, accuracy, and Positive Predictive Value (PPV).

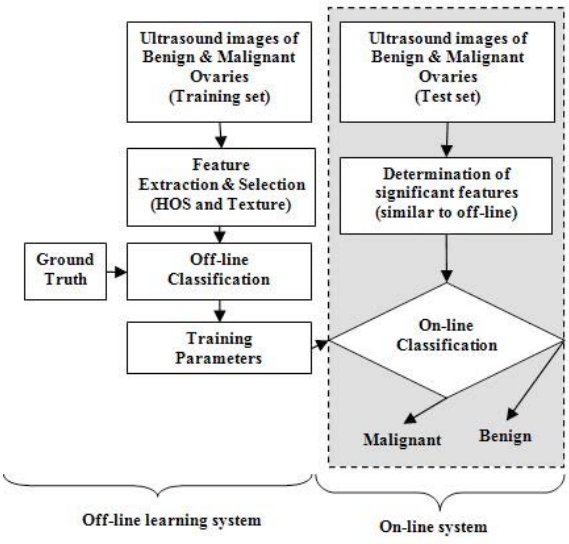


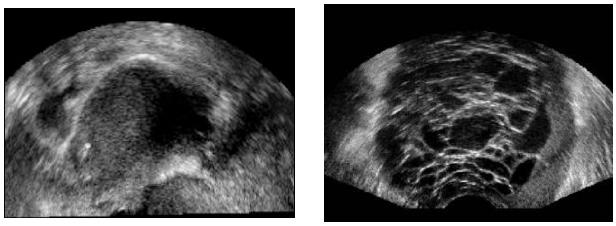
Fig. 1. Block diagram of the proposed system; The blocks outside the dotted shaded rectangular box represent the flow of off-line training system, and the blocks within the dotted box represent the on-line real-time system

B. Data

Twenty women (age: 29-74 years; mean ± SD = 49.5±13.48) were consecutively selected during pre-surgical evaluation. The study was approved by the Institutional Review Board. Informed consent was obtained. Patients with no anatomopathological evaluation were excluded from the study. Biopsies indicated that 10 had malignancy in their ovaries and 10 had benign conditions. All patients were evaluated by 3D-transvaginal ultrasonography using a Voluson-I (GE Medical Systems) according to a predefined scanning protocol using 6.5 MHz probe frequency A 3D volume of the whole ovary was obtained. Volume acquisition time ranged from two to six seconds depending on the size of the volume box. In cases where a given adnexal mass contained more than one solid area, and hence, had more than one volume stored, only the volume best visualizing the mass was chosen for further analysis. Fig. 2 shows typical ultrasound images of benign and malignant classes. We chose the middle 100 images from each volume from each subject. Thus, the evaluated database consisted of 1000 benign images and 1000 malignant images.

C. Feature Extraction

Usually the histopathologic cytoarchitecture of malignant tumors is different from benign neoplasm with several areas having intra-tumoral necrosis.[17,18] These changes manifest as non-linear changes in the texture of the acquired ultrasound images which are quantified by texture-based and HOS based features.



(B1)
Histology: Endometrioma; Echo: characteristic diffuse, low-level echoes of the endometrioma giving a solid appearance.

(M1)
Histology: Borderline malignant tumor; Echo: multiloculate echo-pattern with multiple thick septa.

Fig. 2. Ultrasound images of the ovary: (B1) Benign condition (M1) Malignant tumor

Texture Features:

Fractal Dimension: Fractal Dimension (FD) indicates the irregularity in the pixel intensities of the image. Consider a surface *S* in Euclidean n-space. This surface is self-similar if it is the union of N_r non-overlapping copies of itself scaled up or down by a factor of *r*. In this work, we used the modified differential box counting with sequential algorithm to calculate FD.[19] The input of the algorithm is the gray-scale image where the grid size is in the power of 2 for efficient computation. Maximum and minimum intensities for each (2 x 2) box are obtained to sum their difference, which gives the *M* and *r* by *s/M* where *M=min(R,C)*, *s* is the scale factor, *R* and *C* are the number of rows and columns, respectively. When the grid size gets doubled, *R* and C reduce to half of their original value and above procedure is repeated iteratively until *max(R,C)* is greater than 2. Linear regression model is used to fit the line from plot $log(N_r)$ vs. *log(1/r)* and the slope gives the *FD*.

Gray level Co-occurrence Matrix (GLCM): The elements of the GLCM $C_d(i,j)$ are made up of the relative number of times the gray level pair *(a, b)* occurs when pixels are separated by the distance *(a,b)* = (1, 0).[20] 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

$$P_d(i,j) = \frac{C_d(i,j)}{\sum_i \sum_j C_d(i,j)} \tag{1}$$

Based on Eqn. (1), we obtain the correlation feature, which is a measure of image linearity:

$$Correlation = \frac{\sum_i \sum_j [ij P_d[i,j]] - \mu_i \mu_j}{\sigma_i \sigma_j} \tag{2}$$

where

$$\mu_i = \sum i P_d[i,j], \qquad \sigma_i^2 = \sum i^2 P_d[i,j] - \mu_i^2 \tag{3}$$

$$\mu_j = \sum j P_d[i,j], \qquad \sigma_j^2 = \sum j^2 P_d[i,j] - \mu_j^2 \tag{4}$$

HOS Features:

Higher order statistics denote higher order moments (order greater than two) and non-linear combinations of higher order moments, called the higher order cumulants. They help to extract information on the phase and nonlinearities present in the signal.[21] Prior to the extraction

of HOS based features, the pre-processed images are first subjected to Radon transform to convert a 2D image into a 1D signal at various angles.[22] This 1D signal is then used to determined the bispectrum, $B(f_1,f_2)$ which is a complex valued product of three Fourier coefficients given by

$$B(f_1,f_2) = E[A(f_1)A(f_2)A^*(f_1+f_2)] \quad (5)$$

where $A(f)$ is the Fourier transform of a segment of a single realization of the random signal $a(nT)$, n is an integer index, T is the sampling interval and $E[.]$ stands for the expectation operation. $A^*(f_1+f_2)$ is the conjugate at frequency (f_1+f_2). The function exhibits symmetry, and is computed in a non-redundant/ principal domain region Ω .[21] The bispectrum phase entropy[23] is defined as:

Bispectrum Phase Entropy:

$$ePRes = \sum_n p(\psi_n) logp(\psi_n) \quad (6)$$

where

$$p(\psi_n) = \frac{1}{L}\sum_{\Omega} I(\phi(B(f_1,f_2)) \in \psi_n) \quad (7)$$

$$\psi_n = \{\phi \mid -\pi + 2\pi n/N \le \phi < -\pi + 2\pi(n+1)/N\}, \quad n = 0,1,...,N-1 \quad (8)$$

where L is the number of points within the region Ω, ϕ is the phase angle of the bispectrum, and $I(.)$ is an indicator function which gives a value of 1 when the phase angle is within the range depicted by ψ_n in Eqn. (8). We also derived three bispectrum entropies that are defined below[24] in order to quantify the non-linear changes in the ultrasound images.

Normalized Bispectral Entropy (*e1Res*):

$$e1Res = -\sum_i p_i logp_i \quad (9)$$

where

$$p_i = \frac{|B(f_1,f_2)|^{index}}{\sum_{\Omega}|B(f_1,f_2)|^{index}} \quad (10)$$

where index= 1 for *e1Res*, 2 for Normalized Bispectral Squared Entropy (*e2Res*), and 3 for Normalized Bispectral Cubed Entropy (*e3Res*). We extracted normalized bispectrum entropy, normalized bispectral squared entropy, and normalized bispectral cubed entropy for every one degree of Radon Transform between 0 to 180 degrees. Thus, the total number of extracted features would be 724 (181 x 4).

D. Decision Tree (DT) classifier

In the case of Decision Trees (DT), the input features are used to construct a tree, and then a set of rules for the different classes are derived from the tree.[25] These rules are used for determining the class of an incoming new image.

III. Results

A. Selected Features

After the feature extraction process, there were a total of 726 features (1 FD, 1 texture based, and 724 HOS based). Using all these feature would given rise to curse of dimensionality problem[22] and over-fitting of classifiers. Therefore, we used Student's *t-test*[26] to select significant features. Table I presents the Mean ± Standard Deviation (SD) values of the selected features for both the benign and malignant classes. The low *p*-value indicates that listed four features, namely, FD, Correlation, Bispectral phase entropy at Radon Transform angles of 170º and 171º are significant.

TABLE I
SIGNIFICANT FEATURES THAT HAD A P-VALUE < 0.0001 AND THEIR RANGES (MEAN ± STANDARD DEVIATION) FOR BENIGN AND MALIGNANT CLASSES.

Feature	Benign	Malignant	*p*-value
FD	2.25±4.12E-02	2.26±2.05E-02	< 0.0001
Correlation	0.98±9.01E-03	0.97±8.50E-03	< 0.0001
ePRes (170º)	2.96±0.47	2.91±0.52	< 0.0001
ePRes (171º)	2.99±0.47	2.92±0.52	< 0.0001

B. Classification Results

Ten-fold stratified cross validation was used to evaluate these classifiers. In this method, the dataset was split into ten parts, each part containing the same proportion of images from both classes. In the first fold, nine parts were used for training and the tenth part was used for testing and for calculation of the performance measures. This protocol was repeated nine more times with a different part as the test set. The averages of the performance measures (sensitivity, specificity, Positive Predictive Value (PPV), and accuracy) obtained during the testing phase of each fold are reported as the final performance measures for that classifier. The averages of the performance measures obtained in the 10 folds are reported in Table II. It is evident that the simple decision tree classifier presented the high accuracy of 95.1%, sensitivity of 92.5% and specificity of 97.7%. The advantage of the DT classifier is that this classifier uses rules to classify a new image. These rules are comprehensible to the end-user, and hence, allow the physicians to more confidently accept the result from the classifier. This is not the case with classifiers such as the neural networks which, in most cases, are not transparent in the way in which they determine the class label.

TABLE II
CLASSIFICATION RESULTS OBTAINED USING THE DT CLASSIFIER
A: ACCURACY; SN: SENSITIVITY; SP: SPECIFICITY

SVM	A (%)	PPV (%)	Sn (%)	Sp (%)
Linear	95.1	97.8	92.5	97.7

IV. Discussion

Renz *et al.*[3] used age and results of 30 blood tests as features which resulted in an accuracy of 92.9%. Among the mass spectra based studies,[4-8] the reported accuracies ranged from 84% to 100%. Even though the accuracies are high, the use of these techniques is limited by the availability and cost of the necessary equipment for data analysis. Tailor *et al.*[9] and Biagiotti *et al.*[11] both used several demographic and

quantities from ultrasonography images to obtain classification sensitivity of 100% and 96%, respectively. Bruning *et al.*[10] developed a knowledge-based system called ADNEXPERT that used histopathologic and sonographic data for diagnosis of adnexal tumors and achieved an accuracy of 71%. All these three studies[9-11] used features based on evaluations made by the operator, and hence, these features may be subjective. Zimmer *et al.* [12] used grey-level intensity variations from a segmented region of interest, and classified the tumour into three main categories (cyst, solid and semi-solid), and obtained a low accuracy of 70% for tumours containing solid portions. Lucidarme *et al.*[13] used the Ovarian HistoScanning technique and reported 91.73% accuracy.

In this work, we have proposed a CAD technique which uses a novel combination of four features in the DT classifier to output a high accuracy of 95.1%. The proposed system uses the whole ultrasound image, automatically extracts features and uses them in the DT classifier to predict the class of the patient. Only four easily determinable features have been used which reduces the computational load and time. On the limitations side, for medical legal concerns, the radiologists have to store all CAD findings and images which increase digital storage requirements. We also believe that there is more room for improvement in accuracy. Therefore, in future, we intend to analyze other texture features to determine more discriminating features. Moreover, the clinical applicability of our proposed technique has to be established with more studies containing larger image databases from multi-ethnic groups. We also intend to extend the study to 3D, wherein we would include the spatial information of the 3D slices taken from each patient for analysis.

V. Conclusion

In this work, we determined a novel combination of four texture and HOS based features that adequately quantify the non-linear changes in both benign and malignant ovarian ultrasound images. These features were used in a Decision Tree classifier which registered a high accuracy of 95.1%, sensitivity of 92.5% and specificity of 97.7%. The developed classifier is robust as it was evaluated with 1000 benign and 1000 malignant samples using 10-fold stratified cross validation. Even though the presented preliminary results are based on images from only 20 cases, the highly discriminating features, cross-validation technique used, and the resulting good accuracy make this technique a promising one. The CAD tool would be a more objective alternative to manual analysis of ultrasound images which might result in inter-observer variations.

References

[1] R. C. Bast, D. Badgwell, Z. Lu, et al., "New tumor markers: CA125 and beyond," *Int. J. Gynecol. Cancer*, vol. 15, pp. 274-281, 2005.

[2] S. I. Zaidi, " Fifty years of progress in gynecologic ultrasound," *Int. J. Gynaecol. Obstet.*, vol. 99, pp. 195-197, 2007.

[3] C. Renz, J. C. Rajapakse, K. Razvi, and S. K. C. Liang, "Ovarian cancer classification with missing data," *Proc. 9th Intl. Conf. Neural Information Processing*, vol. 2, pp. 809-813, 2002.

[4] A. Assareh and M. H. Moradi, "Extracting efficient fuzzy if-then rules from mass spectra of blood samples to early diagnosis of ovarian cancer," *IEEE Symposium on Computational Intelligence and Bioinformatics and Computational Biology*, pp. 502-506, 2007.

[5] T. Z. Tan, C. Quek, G. S. Ng, and K. Razvi, "Ovarian cancer diagnosis with complementary learning fuzzy neural network," *Artif. Intell. Med.* , vol. 43, pp. 207-222, 2008.

[6] H. Meng, W. Hong, J. Song, and L. Wang, "Feature extraction and analysis of ovarian cancer proteomic mass spectra," *2nd Intl. Conf. on Bioinformatics and Biomedical Engineering*, pp. 668-671, 2008.

[7] K. L. Tang, T. H. Li, W. W. Xiong, and K. Chen, "Ovarian cancer classification based on dimensionality reduction for SELDI-TOF data," *BMC Bioinformatics*, vol. 11, pp. 109, 2010.

[8] F. Petricoin, "Use of proteomic patterns serum to identify ovarian cancer," *The Lancet*, vol. 359, pp. 572-577, 2002.

[9] A. Tailor, D. Jurkovic, T. H. Bourne, W. P. Collins, and S. Campbell, "Sonographic prediction of malignancy in adnexal masses using an artificial neural network" *Br. J. Obstet. Gynaecol.,* vol. 106, pp. 21-30, 1999.

[10] J. Brüning, R. Becker, M. Entezami, et al., "Knowledge-based system ADNEXPERT to assist the sonographic diagnosis of adnexal tumors," *Methods Inf. Med.* vol. 36, pp. 201-206, 1997.

[11] R. Biagiotti, C. Desii, E. Vanzi, and G. Gacci, "Predicting Ovarian Malignancy: Application of Artificial Neural Networks to Transvaginal and Color Doppler Flow US," *Radiology*, vol. 210, pp. 399-403, 1999.

[12] Y. Zimmer, R. Tepper, and S. Akselrod, "An automatic approach for morphological analysis and malignancy evaluation of ovarian masses using B-scans," *Ultrasound Med. Biol.* vol. 29, pp. 1561–1570, 2003.

[13] O. Lucidarme, J. P. Akakpo, S. Granberg, et al., "A new computer-aided diagnostic tool for non-invasive characterisation of malignant ovarian masses: results of a multicentre validation study," *Eur. Radiol.*, vol. 20, pp. 1822-1830, 2010.

[14] R. E. Bellman, *Dynamic Programming*. Courier Dover Publications, 2003.

[15] T. Hata, T. Yanagihara, K. Hayashi, et al., "Three-dimensional ultrasonographic evaluation of ovarian tumours: a preliminary study," *Hum. Reprod.*, vol. 14, pp. 858-861, 1999.

[16] M. Laban, H. Metawee, A. Elyan, M. Kamal, M. Kamel, and G. Mansour, "Three-dimensional ultrasound and three-dimensional power Doppler in the assessment of ovarian tumors," *Int. J. Gynaecol. Obstet.*, vol. 99, pp. 201-205, 2007.

[17] S. Guerriero, J. L. Alcazar, M. A. Pascual, et al., "Intraobserver and interobserver agreement of grayscale typical ultrasonographic patterns for the diagnosis of ovarian cancer," *Ultrasound Med. Biol.,* vol. 34, pp. 1711-1716, 2008.

[18] A. C. Testa, A. Gaurilcikas, and A. Licameli, et al., "Sonographic features of primary ovarian fibrosarcoma: a report of two cases," *Ultrasound Obstet. Gynecol.*, vol. 33, pp.112-115, 2009.

[19] M. K. Biswas, T. Ghose, S. Guha, and P. K. Biswas, "Fractal dimension estimation for texture images: A parallel approach," *Pattern Recogn Letters*, vol. 19, pp. 309-313, 1998.

[20] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man. Cybern.* SMC-3, pp.610–621, 1973

[21] C. Nikias and A. Petropulu, *Higher-Order Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[22] A. Ramm and A. Katsevich, *The radon transform and local tomography*. CRC Press, 1996.

[23] K. Chua, V. Chandran, U. Acharya, and C. Lim, "Application of higher order spectra to identify epileptic EEG," *J. Med. Syst.,* pp. 1–9, 2010, 10.1007/s10916-010-9433-z.

[24] K. Chua, V. Chandran, U. Acharya, and C. Lim, "Analysis of epileptic EEG signals using higher order spectra," *J. Med. Eng. Technol.*, vol. 33, pp. 42-50, 2009.

[25] D. T. Larose, Decision Trees. In: *Discovering Knowledge in Data: An introduction to data mining*. New Jersey, USA: Wiley Interscience, pp. 108-126, 2004.

[26] J. F. Box, "Guinness, gosset, fisher, and small samples," *Statist. Sci.*,vol. 2, pp. 45–52, 1987.