IMPROVING DIAGNOSTIC ACCURACY IN THE CLASSIFICATION OF THYROID CANCER BY COMBINING QUANTITATIVE INFORMATION EXTRACTED FROM BOTH ULTRASOUND AND CYTOLOGICAL IMAGES

Stavros Tsantis†, Dimitris Glotsos†, Panagiota Spyridonos†, Giannis Kalatzis++, Nikos Dimitropoulos+++,
George Nikiforidis†, Dionisis Cavouras++
† Medical image processing and analysis laboratory
Department of Medical Physics, University of Patras
26500 Patras, Greece
e-mail: tsantis@med.upatras.gr cavouras@teiath.gr

++ Department of Medical Instrumentation Technology
Technological Institute of Athens
12 210 Aigaleo, Athens, Greece

+++ Department of Medical Imaging
EUROMEDICA Medical Center
11 527 Athens, Greece

++++ Department of Pathology
University Hospital of Patras
26500 Patras, Greece

Keywords: classification, microscopic images, ultrasound images, thyroid, support vector machines.

Abstract. Purpose: Develop a computer based image analysis system for the automatic characterization of thyroid nodule cancer as low or high risk via the quantitative analysis of cytological and sonographic images.

Materials and methods: Clinical data comprised 120 pairs of sonographic and cytological images, with every pair corresponding to a distinct patient. Ultrasound images were processed for de-noising and segmentation to optimally highlight the cancerous region from the remaining background. From this region, 20 textural features were generated. Cytological images were segmented to separate nuclei from background. 20 morphological and textural features were extracted from segmented nuclei. An SVM classifier was designed to discriminate low from high-risk cases evaluating the combined set of 40 ultrasound and cytological features. For comparative evaluation a Bayesian classifier was also applied to the same data set. Results: SVM achieved 93.3% precision in distinguishing correctly low from high-risk thyroid nodules. The corresponding performance of the Bayesian classifier was 90.8%. An exhaustive search algorithm determined the optimal feature combination comprising 2 ultrasonic and 1 cytological textural features. Conclusions: The combined analysis of cytological and ultrasound images of thyroid nodules may be indicative of thyroid nodule’s malignancy-risk and may be of value to patient management.

1 INTRODUCTION

Recent studies have indicated that the combination of ultrasonography and cytology improves the accuracy of diagnosis of thyroid nodules [1]. Ultrasound scans solely, cannot provide a definite diagnosis of thyroid cancer, generating the necessity to be combined with Fine Needle Aspiration (FNA) biopsy examination. In order to conclude on the degree of tumour abnormality, doctors initially examine ultrasound images and subsequently investigate FNA cytological material under the microscope [2]. Based on the nuclei appearance, thyroid nodules are then, classified as low or high risk according to the WHO (World Health Organization) scheme. It is critical to accurate discriminate low from high risk tumours, because are treated differently [3].

Nevertheless, it is common even for biopsy results to be non-diagnostic, either due to limitations in the number of cells obtained, or due to difficulty in making an accurate diagnosis for some types of thyroid pathology [4]. Additionally, the subjectivity in the interpretation by the pathologists has been shown to provoke diagnostic errors [5]. The stakes are heightened considering that errors result in false treatment planning with adverse effect in patient survival.

To overcome these limitations, computer based systems were introduced to assist pathologists to more...
objective decisions [6-9]. These studies though, have exclusively concerned either computer-based systems examining only ultrasound images or only FNA cytological images. In this work we propose a computer based system that automatically characterizes thyroid nodule cancer as low or high risk by combining quantitative information generated from both ultrasound and FNA image analysis.

2 MATERIALS AND METHODS

2.1 Clinical data

Clinical data comprised 120 pairs of sonographic and cytological images, with every pair corresponding to a distinct patient. Material was obtained by the Department of medical imaging of the EUROMEDICA private clinic.

2.2 Ultrasound image processing and analysis

The study comprised 120 ultrasonic (US) images displaying thyroid nodules of 120 patients. All US examinations were performed on an HDI-3000 ATL digital ultrasound system with a wide band (5-12 MHz) probe using various scanning methods such as longitudinal, transversal and sagittal cross sections of the thyroid gland. Each US image was digitized via connecting the video output of the ultrasound scanner to a Screen Machine II frame grabber using 512×512×8 image resolution. A US image of thyroid nodule is presented in figure 1. All thyroid nodules were pathology proven cases. They were separated by the experienced physician (N.D.) into two classes; 78 images were assigned to the low-risk and 42 to the high-risk group. Class assignments were also based on the following US characteristics [10]. The low risk class comprised iso-echoic or hyper-echoic solid nodules with or without cystic change and coarse calcification, and the high-risk class contained hypo-echoic solid nodules with regular borders or cystic nodules with solid components. The boundary of each nodule was delineated by the physician employing an easy-to-use interactive software program implemented in C++ for the purposes of the present study. Data processing was performed on a Pentium IV, 2.4 GHz computer. A number of textural features were automatically calculated from the segmented ROI of each thyroid nodule. Textural features are related to the gray-tone structure of the thyroid nodule as depicted on the ultrasound unit and carry information relevant to the risk factor of malignancy. Four features were computed from the nodule’s gray-tone histogram, 10 from the co-occurrence matrix [11] and 6 from the run-length matrix [12].

Figure 1. Typical US image of thyroid nodule

2.3 FNA cytological image processing and analysis

From each biopsy, a pathologist specified the most representative region. From this region images were captured (768x576x8 bit) using a using a light Zeiss Axiostar plus microscope connected to a Leica DC 300 F color video camera. A typical FNA image of thyroid cancer is illustrated in figure 2. Nuclei were automatically segmented from background. The segmentation algorithm is described elsewhere [13]. A set of 17 textural and 3 morphological features were generated to encode tumor malignancy for each case (patient). Textural features were the mean value, standard deviation, skewness and kurtosis extracted from nuclei histogram and 13 descriptors created based on the co-occurrence and run length matrix [11-12]. Morphological features comprised measurements of nuclear area, concavity and roundness.
2.4 Design of the classification scheme

An SVM classifier [14] was designed to discriminate low from high-risk cases evaluating the combined set of 40 ultrasound and cytological features. An exhaustive search in feature space was utilized to extract the best feature vector combination. According to this method [15], features were combined in all possible ways (combinations of 2, 3, 4, 5 and 6). Each combination was evaluated using the leave-one-out method [16]. The best combination selected, was the one that resulted the smallest classification error with the minimum number of support vectors. For comparative evaluation, a Bayesian classifier [17] was applied to the same data. Results are indicated in tables 1 and 2.

2.5 Support vector machines

SVM are state of the art algorithms, originated from statistical learning theory. To apply SVM for binary classification problem the machine initially maps the input space through a transformation function and subsequently define the hyperplane that has maximum distance from its closest training data [18].

The discriminant function for two class separation problems is given in equation 1.

\[
g(x) = \text{sign} \left( \sum_{i=1}^{N} \alpha_i y_i K(x, x_i) + b \right)
\]

where \(x_i\) training data belonging to either class \(y_i \in \{+1,-1\}\), \(N\) the number of training samples, \(\alpha_i, b\) weight coefficients and \(K\) the kernel function. Adjustable parameter is the cost factor \(C\) that specifies the importance of misclassifications. In this work \(C\) was experimentally determined as equal to 100. In the present work, the SVM classifier was designed employing various polynomial kernels up to the 4th degree and the radial basis function (RBF) kernel. These kernel functions are described by the following relations (eq. 2-3):

\[
k_{\text{POLYNOMIAL}}(x_1, x_2) = \left( (x_1^T x_2) + 1 \right)^d, \; d = \text{degree}
\]

\[
k_{\text{RBF}}(x_1, x_2) = \exp \left( -\frac{\|x_1 - x_2\|^2}{2\sigma^2} \right), \; \sigma = \text{standard deviation}
\]

2.6 Quadratic Bayesian classifier

The Bayes decision theory develops a probabilistic approach to pattern recognition, based on the statistical nature of the generated features. The Bayes discriminant function [17] for class \(i\) and for pattern vector \(x\) is given by:

\[
g(x) = \ln P_i - \frac{1}{2} \ln |C_i| - \frac{1}{2} [(x - m_i)^T C_i^{-1} (x - m_i)]
\]

where \(P_i\) is the probability of occurrence of class \(i\), \(m_i\) is the mean feature vector of class \(i\), and \(C_i\) is the covariance matrix of class \(i\).
For the SVM-classifier employing the polynomial kernel of 2nd degree, best feature combination comprised
the mean gray-level value of the thyroid ROI’s histogram the sum variance from the co-occurrence matrix from
the US features and entropy from the cytological features [10-11].

The mean value was calculated by (eq. 5):

$$\mu_g = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} I(x, y)$$

where $I(x,y)$ is the gray-tone value at coordinates $(x,y)$ of the image matrix of dimensions $N_x \times N_y$.

The sum variance and entropy were determined by (eq. 6-7):

$$SVA = \sum_{i=2}^{2N_g} \left[ i - \sum_{i=2}^{2N_g} \left( p_{x+y}(i) \right)^2 \right] p_{x+y}(i),$$

$$ENT = -\sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} p(i,j) \log(p(i,j))$$

where $N_g$ is the number of distinct gray levels in the image and $p_{x+y}(k)$ is given by the following relation (eq. 8)

$$p_{x+y}(k) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} p(i,j), \quad k = 2, 3, ..., 2N_g$$

where $p(i,j)$ is the normalized entry of the co-occurrence matrix averaged over the four directions ($0^\circ$, $45^\circ$, $90^\circ$, $135^\circ$) [10].

3 RESULTS AND DISCUSSION

With the SVM classifier low from high risk cases were discriminated with an accuracy of 93.3%. Optimum
parameter setting were $C=100$, and polynomial kernel of degree 2. The number of support vectors was 13. For
the same data, the Bayesian classifier resulted in 90.8% accuracy.

Table I gives a detailed account of the SVM-2nd degree polynomial kernel classification accuracies obtained
by the LOO method, employing the mean gray-level, sum variance and entropy features combination. Seventy-
three of the low-risk thyroid nodules were correctly classified while five nodules were incorrectly assigned to
the high-risk class, giving a classification accuracy of 93.6%. In the case of the high-risk thyroid nodules, thirty-
nine were assigned to the correct class while only three were wrongly classified to the low-risk class, scoring
92.7% class discrimination accuracy. Overall, the SVM achieved 93.3% precision in distinguishing correctly
low-risk from high-risk thyroid nodules with 13 support vector. Figure 3 shows a 3-dimensional scatter
diagrams of the mean gray-level, sum variance, and entropy, as well the decision boundary drawn by the SVM
classifier respectively.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>System classification</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk</td>
<td>Low risk</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>High risk</td>
<td>5</td>
</tr>
<tr>
<td>High risk</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>39</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Truth Table demonstrating classification results in the task of separating low from high-risk cases of
thyroid nodule cancer using the SVM classifier ($C=100$, polynomial kernel with $d=2$)
The maximum classification accuracy corresponds to the minimum number of support vectors involved in the design of the SVM-classifier. The number of support vectors for the best feature combination corresponds approximately with the 12% of the number of training points. The small number of support vectors is indicative of the SVM manageable class separability.

Table 2 is the truth table giving the classification performance of the Bayesian classifier using the best feature combination (mean gray-level value, sum variance, and entropy). Seventy-three of the low-risk and 36 of the high-risk thyroid nodules were correctly classified using the LOO method, resulting in group classification accuracies of 93.6% and 85.7% respectively and overall precision of 90.8%. Figure 4 shows a 3-dimensional scatter diagrams of the mean gray-level, sum variance, and entropy, as well the decision boundaries drawn by the Quadratic Bayesian classifiers respectively.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Low risk</th>
<th>High risk</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk</td>
<td>73</td>
<td>5</td>
<td>93.6%</td>
</tr>
<tr>
<td>High risk</td>
<td>6</td>
<td>36</td>
<td>85.7%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td></td>
<td></td>
<td>90.8%</td>
</tr>
</tbody>
</table>

Table 2. Truth Table demonstrating classification results in the task of separating low from high-risk cases of thyroid nodule cancer using the Bayesian classifier.
Echogenicity and the existence of different structures inside the thyroid nodule have been indicated as important factors leading to thyroid malignancy in US images [...]. Entropy describes the smoothness of chromatin within nuclei. It takes low values for low risk nuclei, since these are smoother and high values for high risk nuclei due to the coarseness of their texture. We have developed a quantitative method for assessing the malignancy risk factor of thyroid nodules combining features from both US and cytology images.

Comparing the SVM with the Quadratic Bayesian classifier it is evident that the latter showed excellent performance, however without reaching the SVM's precision. This is indicative of the effectiveness of the SVM. The penalty, however, that had to be paid for employing the SVM algorithm was much higher processing time during classifier design (training).

4 CONCLUSIONS

In conclusion, an efficient classification system was designed, based on the SVM algorithm, for assessing the malignancy risk factor of thyroid nodules from US and cytology images. This system could be a useful diagnostic tool by providing a second opinion to the physician and, thus, it may be of value to patient management.

5 ACKNOWLEDGEMENTS

The present research was carried for the project "Computer-based system for the automatic diagnosis of thyroid nodule cancer" co-funded by 75% from the European Union and 25% from the Greek Government under the framework of the Education and Initial Vocational Training Program - Archimedes.

6 REFERENCES


