PROD. TYPE: COM
PP: 1-8 (col.fig.: nil)
CBM855 MODE+
ED: Shija
PAGN: Jagan – SCAN: global



ELSEVIER

Computers in Biology and Medicine ▮▮▮ (▮▮▮▮) ▮▮▮–▮▮▮

Computers in Biology and Medicine

www.intl.elsevierhealth.com/journals/cobm

# Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological images

Antonis Daskalakis[a,*], Spiros Kostopoulos[a], Panagiota Spyridonos[a], Dimitris Glotsos[a], Panagiota Ravazoula[b], Maria Kardari[b], Ioannis Kalatzis[c], Dionisis Cavouras[c], George Nikiforidis[a]

[a]*Medical Image Processing and Analysis Group, Lab of Medical Physics, School of Medicine, University of Patras, Rio, Patras 265 04, Greece*
[b]*Department of Pathology, University Hospital, Rio, Patras 265 04, Greece*
[c]*Medical Signal and Image Processing Lab, Department of Medical Instruments Technology, Technological Educational Institute of Athens, Ag. Spyridonos Street, Aigaleo, 122 10 Athens, Greece*

Received 1 December 2006; accepted 28 September 2007

**Abstract**

A multi-classifier diagnostic system was designed for distinguishing between benign and malignant thyroid nodules from routinely taken (FNA, H&E-stained) cytological images. To construct the multi-classifier system, several combination rules and different mixtures of ensemble classifier members, employing morphological and textural nuclear features, were comparatively evaluated. Experimental results illustrated that the classifier combination k-NN/PNN/Bayesian and the majority vote rule enhanced significantly classification accuracy (95.7%) as compared to best single classifier (PNN: 89.6%). The proposed system was designed with purpose to be utilized in daily clinical practice as a second opinion tool to support cytopathologists' decisions, when a definite diagnosis is difficult to be obtained.


*Keywords:* Multi-classifier systems; Quantitative analysis of cell nuclei; Computer-assisted microscopy; Hematoxylin & Eosin; Cytological images; Thyroid nodules

## 1. Introduction

Nodular thyroid lesions can be either benign or malignant. Malignant nodules often lead to cancer-related deaths and their aggressiveness increases significantly with patients' age [1]. Physical examination combined with laboratory tests and ultrasonography are primarily used for the initial detection of thyroid nodules. Although these techniques are useful in detecting the extent and composition (cystic or solid) of thyroid masses [2], they cannot provide a definite diagnostic differentiation between benign and malignant nodules [3].

Fine needle aspiration (FNA) biopsy examination is currently considered as the test of choice for the diagnosis and management of thyroid nodules with an accuracy approaching 95% [1]. However, routine screening of FNA microscopic slides is an elaborating and time consuming task, even for skillful cytopathologists, since it requires the visual inspection and evaluation of subtle nuclei morphological and textural differentiations [4,5].

In order to objectify the diagnostic process, computer-assisted microscopy systems have been introduced [3,6–11]. These systems have been designed to investigate whether morphological and textural features of cell nuclei may be used to quantitatively assess thyroid diseases. Nagashima et al. [12] have reported significant statistical differences in nuclear area and nuclear morphometry between the subtypes of follicular thyroid lesions using Student's *t*-test. Tseleni et al. [10] have investigated the differences in nuclei sizes between various types of thyroid carcinomas. Galera-Davidson et al. [13] have found that optical density values were significantly different in malignant and normal nuclei following stepwise discriminant analysis. Ferrer-Roca et al. [3,14] have attempted to separate normal thyroids from adenomas and from carcinomas using densitometric, morphometric, and flow cytometric parameters and Fisher linear discriminant analysis. Although quantitative

* Corresponding author. Tel.: +30 2610 99745; fax: +30 2610 997882.
*E-mail address:* daskalakis@med.upatras.gr (A. Daskalakis).

0010-4825/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2007.09.005



Please cite this article as: A. Daskalakis, et al., Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological..., Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2007.09.005

cell nuclei morphometry carries information for classifying individual cells, its usefulness in classifying patient cases has been controversial [3,13,15]. More recent studies have investigated the potential of pattern recognition techniques (back propagation neural networks and non-parametric classifiers) for improving the diagnostic accuracy of thyroid lesions in classifying patient cases [8,9]. According to these approaches, a case is classified as malignant if a certain proportion of the cell nuclei are labeled as malignant by the system, employing the May Grunvald Giemsa specialized staining protocol.

In the present study, we introduce a multi-classifier system to address the problem of diagnosing thyroid diseases, however, using cytological images of inferior quality, produced by a routinely used in most cytological laboratories staining protocol (Hematoxylin & Eosin (H&E) [16]). The challenge of using multi-classifiers systems to improve medical diagnosis has been investigated by several researchers but for other applications [17–20]. Peng has suggested a support vector machine (SVM)-based classifier combination scheme for classification of microarray data. Hayashia et al. have shown the application of NN-multi-classifier systems for diagnosing hepatobiliary disorders. Sboner et al. [19] have used an ensemble of linear discriminant analysis, k-NN and decision tree for improving the diagnosis of early melanoma. Lum et al. [20] have investigated several classifier combination schemes, employing the SVM and the Bayesian classifiers in detecting bone fractures.

Our motivation for exploring various classifier-combinations was to investigate whether by employing such structures instead of single classifiers we could improve the discrimination accuracy between benign and malignant thyroid nodules using low quality H&E-stained images. In this way, it might be fully exploited by cytopathologists in daily clinical practice in order to minimize the risk of false patient selection for surgical operation, while detecting malignancy with a high accuracy.

## 2. Material and methods

### 2.1. Material

Clinical material comprises 115 FNA biopsies of thyroid nodules of 115 patients. FNA biopsies were collected from the Department of Pathology of the University Hospital of Patras, Greece. Patients were pre-selected in order to safeguard an adequate data sample for malignant and benign cases. Biopsies were stained with H&E and were characterized as benign (53/115) or malignant (62/115), according to the WHO criteria by two experienced physicians, a cytopathologist (M.K.) and a histopathologist (P.R.). Malignant thyroid nodules were pathologically confirmed by histological findings. From each biopsy, the cytopathologist specified the most representative region and marked it on the microscopy slide. From this region, images ($1300 \times 1030 \times 8$ bit) were acquired using a light Zeiss Axiostar-Plus microscope (ZEISS; Germany) connected to a Leica DC 300 F color video camera (LEICA; Germany). Consequently, 115 images were selected corresponding to the 115 patients of the study (one image per patient).

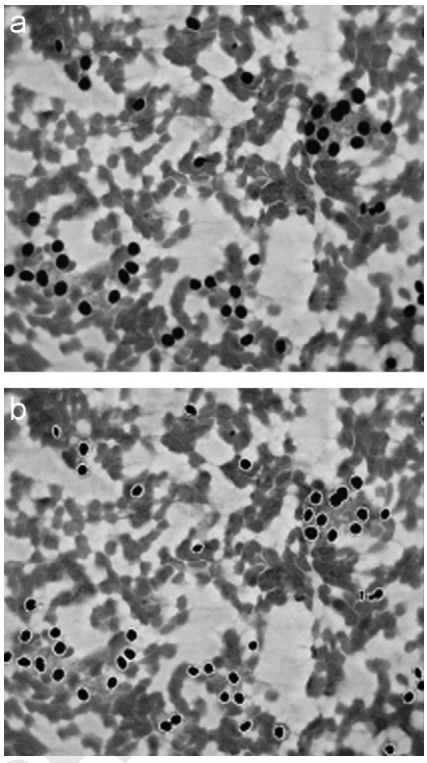


Fig. 1. (a) Original image sample of benign thyroid nodule (magnification $\times 400$) and (b) pixel-based segmentation algorithm delineating cell nuclei in (a).

### 2.2. Design of the classification system

The design of the thyroid nodules classification system was performed in three stages: Images were initially segmented to isolate nuclei from surrounding background. Subsequently, features, encoding nodules' nature (benign-malignant), were extracted from the segmented nuclei. Extracted features were then fed into a multi-classifier structure that comprised five different classifiers (linear least squares minimum distance (LSMD) [21], statistical quadratic Bayesian [22], k-nearest neighbor (k-NN) [22], SVM's [23], and probabilistic neural network (PNN) [24]). The system was tuned to identify benign from malignant nodules, based on nuclei differences in texture and morphology.

### 2.3. Segmentation

Accurate segmentation of nuclei is of crucial importance to guarantee correct results in computer-assisted microscopy [25]. To quantify and investigate the important alterations induced on cancerous thyroid nuclei by the disease, a pixel-based segmentation algorithm was employed, aiming at nuclei delineation (see Fig. 1). This technique relies on the principle of classifying each image pixel as belonging to nuclei or surrounding background based on the textural properties of a small neighborhood region around each pixel. The segmentation algorithm has been previously applied to nuclei identification in microscopic images and has been developed by our group [26–28].



Please cite this article as: A. Daskalakis, et al., Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological..., Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2007.09.005

Table 1
Morphological and textural features extracted from cell nuclei

| Nuclear DNA content and chromatin textural-related nuclear features | | | Features describing shape and size |
|---|---|---|---|
| First-order statistical | Second-order statistical | | |
| Based on DNA histogram | Co-occurrence matrix[a] | Run-length matrix[b] | |
| Density ($D$) | Correlation (Cor) | Short run emphasis (SRE) | Area (Area) |
| Standard deviation (StD) | Entropy ($E$) | Long run emphasis (LRE) | Roundness (Round) |
| Kurtosis (Kurt) | Inverse_diff_moment (IDF) | Gray-level non-uniformity (GLNU) | Concavitiy (Con) |
| Skewness (Skew) | Sum_entropy (SE) | Run-length non-uniformity | Perimeter (Perim) |
| | Angular second moment (ASM) | Run percentage (RP) | |
| | Difference_variance (DifVar) | | |
| | Difference_entropy (DifE) | | |
| | Sum_of variance (SumVar) | | |
| | Sum_of average (SumAv) | | |
| | Absolute_value (AbsVal) | | |
| | Variance (Var) | | |
| | Autocorrelation (Autocor) | | |
| | Contrast(Con) | | |

[a]Each feature was computed with inter-pixel distance $d = 1$.
[b]The number of gray-level runs was $R - L = 4$.

### 2.4. Feature extraction

Following segmentation, 26 morphological and textural features were extracted from each image to quantify the diagnostic information of cell nuclei (Table 1); morphological and textural features were extracted from each segmented nuclei and a representative 26-dimensional mean feature vector was formed for each patient. Morphological features describe the size and shape of nuclei and comprised measures of area, perimeter roundness, and concavity. Textural features, that encode the variations in chromatin distribution within nuclei [29] were calculated from first-order (mean, standard deviation, kurtosis, and skewness) and second-order statistics (features emanating from the run-length [30] and co-occurrence matrices [31]).

### 2.5. Features classification

#### 2.5.1. Multi-classifier system

Multi-classifier systems exploit the idea of combining information from multiple sources with ultimate goal to improve classification results, compared to the use of single classifiers [32]. Such systems have been recently introduced to medical research with promising results [19,33]. In this study we applied five different classifiers to the problem of discriminating benign from malignant thyroid nodules and we tested whether their combination might improve classification performance. The classifiers used were two non-parametric (LSMD [21] and Bayesian [22]) and three parametric (PNN [24], k-NN with Euclidean distant metric [22], and SVMs [23]).

LSMD classification is based on minimizing the mean-square error between the training set and arbitrary pre-selected points in an augmented feature space. The discriminant function of LSMD is given by [21]:

$$d_i(x) = \sum_{i=1}^{d} a_{ci} x_i - b_c, \quad (1)$$

where $d$ is the number of features, $x$ is the unknown pattern vector, $a_c$ are weight coefficients, $b_c$ is a threshold parameter, and $x_i$ the unknown pattern vector elements.

The Bayesian is the optimal statistical classifier, designed to give the minimum probability error for data following Gaussian distribution. The discriminant function of the Bayesian classifier is [22]

$$d_i(x) = \ln(P_i) - \tfrac{1}{2}\ln(|C_i|) - \tfrac{1}{2}[(x - m_i)^{\mathrm{T}} C_i^{-1}(x - m_i)], \quad (2)$$

where $P_i$ is the probability of occurrence of each class $i$, $C_i$ is the covariance matrix, and $m_i$ is the mean value of class $i$.

The PNN is a parametric four-layer feedforward neural network classifier. The PNN determines each class probability density function (PDF) by linearly combining the kernel PDF estimation for each training sample separately for a given class. Its discriminant function is given by [24]

$$d_i(x) = \frac{1}{(2\pi)^{d/2}\sigma^d} \frac{1}{N} \sum_{k=1}^{N} \exp\left[-\frac{(x - x_{ik})^{\mathrm{T}}(x - x_{ik})}{2\sigma^2}\right], \quad (3)$$

where parameter $\sigma$ defines *the* spread of the Gaussian activation function, $N$ is the number of pattern vectors, $d$ is the dimensionality of pattern vectors, and $x_{ik}$ is the $k$th pattern vector of class $i$.

The k-NN is a parametric classifier. Unknown patterns are classified using the following procedure: The k-NN that the unknown pattern is closer to are identified and the unknown pattern is classified to class $j$ if the maximum number of $k_j$ belongs to class $j$ [22].

The SVM classifier again is a parametric classifier that seeks the optimal mapping of input space into a higher dimensional feature space, in which data can be considered linearly separable. The mapping is performed using a non-linear transformation function (kernel), with most common choice the Gaussian



Please cite this article as: A. Daskalakis, et al., Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological..., Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2007.09.005

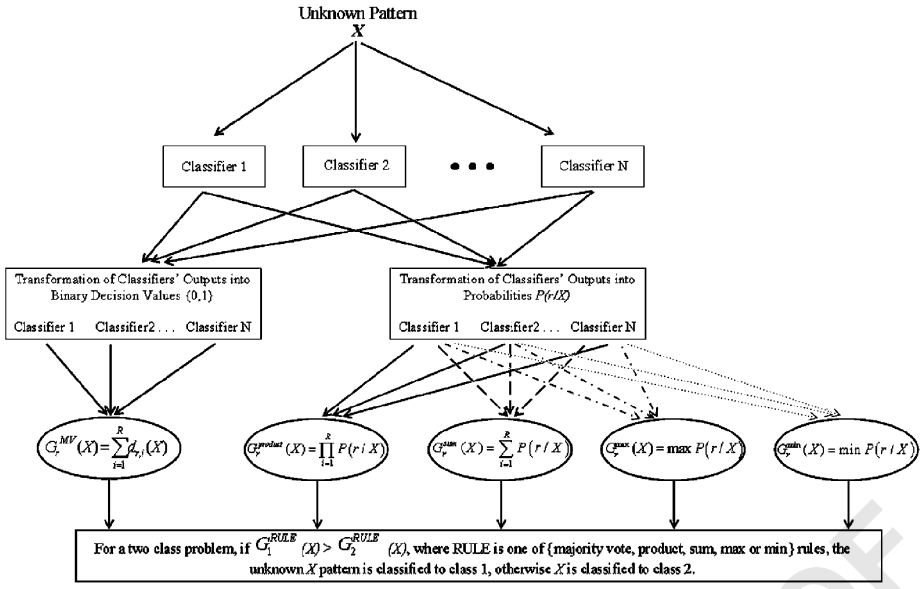


Fig. 2. Different schemas of multi-classifiers schemes.

Radial Basis function:

$$K_{\mathrm{RBF}}(x, x_i) = \exp\left(-\frac{\|z - z_i\|^2}{2\sigma^2}\right), \quad \sigma = \text{spread}. \tag{4}$$

The discriminant function of the SVM classifier for binary classification problems is

$$d(x) = \text{sign}\left(\sum_{i=1}^{N} \alpha_i y_i K_{\mathrm{RBF}}(\mathbf{x}, \mathbf{x}_i) + b\right), \tag{5}$$

where $x_i$ is the training data belonging to either class $y_i \in \{+1, -1\}$, and $\alpha_i$ and $b$ are weight coefficients [23].

The selection of the classifiers was mainly driven from the fact that these algorithms have been individually applied successfully in computer-assisted analysis of cell nuclei [34–36]. Moreover, the k-NN and SVM classifiers have been used as "bench-mark" techniques in many pattern recognition applications [22,23]. Finally, their design is based on different theoretical concepts. This is of particular interest in combining classifiers, since they might explore complementary information for the particular classification task.

A multi-classifier system is valuable if it outperforms best individual classifier's performance [37]. For this reason, individual classifiers were primarily optimized with respect to adjustable parameters settings and functional feature sub-space selection [38]. Adjustable parameters, which were experimentally determined, were the spread of the Gaussian activation function for the PNN classifier, the number of k-neighbors for the k-NN classifier, and the selection of kernel function for the SVM classifier. Feature sub-space selection was performed in order to retain only the feature subset with the highest discrimination information and the least number of features. For this purpose an exhaustive search [22] algorithm was employed by combining features in all possible ways. The classifier's performance for each distinct feature combination was evaluated using the leave one out method [22].

Each individual classifier's performance was optimized in terms of overall classification accuracy i.e. maximize system's sensitivity (probability of a positive test among patients with thyroid malignancy) and specificity (probability of a negative test among patients without thyroid malignancy).

### 2.5.2. Ensemble classifier combination rules

The ensemble combination rules tested in this study were the majority vote rule and the aggregation rules of minimum, maximum, average, and product [37,39,40]. The majority vote scheme goes with the decision when there is a consensus for it or at least more than half of the classifiers agree on it. Consequently, classification outputs from each classifier were combined to form the final classification decision, following a majority vote rule [37], that satisfies

$$G_r(X) = \sum_{i=1}^{R} d_{r,i}(X), \tag{6}$$

where $r$ is the class, $X$ is the unknown pattern vector, $i = 1, 2, \ldots, R$ is the odd number of classifiers involved in the majority vote scheme, $d_{r,j}$ is the binary decision value $\{0, 1\}$, 0 corresponds to misclassification and 1 to correct classification. For a two-class problem, if $G_1(X) > G_2(X)$, the unknown $X$ pattern is classified to class 1, otherwise $X$ is classified to class 2.

Classifier combination using aggregation rules is not straightforward, since individual classifiers' outputs vary due to their different theoretical frameworks (similarity measures, probability measures, etc.). For this reason, individual classifiers outputs were normalized using the min–max normalization technique [41] in order to transform classifiers outputs into probabilities $P(r/X)$, such as the sum probabilities of the classes add up to

Table 2
Best feature subsets for that optimized the performance of individual classifiers

| | LSMD | k-NN | PNN | Bayesian | SVM |
|---|---|---|---|---|---|
| Best feature vector combination | IDF | *D* | *D* | StD | *D* |
| | SRE | Cor | StD | Skew | StD |
| | GLNU | Var | Var | DifE | DifE |
| | Round | Perim | Area | GLNU | Area |

1 1 for each individual classifier ($\sum_{i=1}^{N_c} P(r/X)=1$, where $N_c$ is the number of classes). The aggregation rules for classifying an 3 unknown pattern $X$ to a class $r$ are defined in relations (7–10).

*Product rule*:

$$G_r^{\text{product}}(X) = \prod_{i=1}^{R} P(r/X). \quad (7)$$

*Sum rule*:

$$G_r^{\text{sum}}(X) = \sum_{i=1}^{R} P(r/X). \quad (8)$$

*Max rule*:

$$G_r^{\max}(X) = \max P(r/X). \quad (9)$$

*Min rule*:

$$G_r^{\min}(X) = \min P(r/X), \quad (10)$$

where $P(r/X)$ is the probability of $X$ to belong to class $r$. For 13 a two-class problem, if $G_1^{\text{RULE}}(X) > G_2^{\text{RULE}}(X)$, where RULE is one of {product, sum, max, or min} rules, the unknown $X$ 15 pattern is classified as class 1, otherwise $X$ is classified as class 2. Fig. 2 depicts schemas of multi-classifiers systems utilizing 17 the aforementioned aggregation rules. To ensure the optimum multi-classifier ensemble combination selection an exhaustive 19 search of all possible classifiers combinations was implemented [42]. All algorithms were developed using custom-made Matlab 21 source code.

## 3. Results

23 The individual classifiers were optimized with regard to parameter settings and available feature data. More specifically, 25 for the k-NN classifier, parameter $k$ was experimentally set equal to 5, while for the PNN classifier, the spread of the Gaus- 27 sian function was experimentally set equal to $\sigma = 0.23$. For the SVM classifier, the Gaussian radial basis function kernel 29 yielded the optimal performance with spread set equal to 1. Furthermore, each classifier was optimal designed using its own 31 representation of the input feature space. Best feature vectors and corresponding accuracies for each classifier are given in 33 Tables 2 and 3, respectively.

Choosing the optimum ensemble scheme involved experi- 35 mentation, employing the exhaustive search methodology in ensemble members. Best ensemble scheme was selected as the 37 one that provided highest classification accuracy with minimum

Table 3
Classification performance of each individual classifier in the task of discriminating benign from malignant nodules

| Classifiers | Classification task (%) | | |
|---|---|---|---|
| | Sensitivity | Specificity | Overall accuracy |
| LSMD | 77.4 | 75.5 | 76.5 |
| k-NN | 88.7 | 83.0 | 86.1 |
| PNN | 88.7 | 90.3 | 89.6 |
| Bayesian | 85.5 | 69.8 | 78.3 |
| SVM | 90.3 | 75.5 | 83.5 |

Table 4
Performance rates of several ensemble classifier combination rules, using as ensemble members the Bayesian, PNN, and k-NN classifiers

| Schemes | Ensemble schemes classification (%) | | |
|---|---|---|---|
| | Sensitivity | Specificity | Overall accuracy |
| Minimum | 93.6 | 88.7 | 91.3 |
| Maximum | 85.5 | 77.4 | 81.7 |
| Average | 93.6 | 83.0 | 89.6 |
| Product | 91.9 | 86.8 | 88.7 |
| Majority vote | 96.8 | 94.3 | 95.7 |

number of classifiers involved. Thus, the most effective ensemble scheme comprises the Bayesian, the PNN, and the k-NN 39 classifiers. The latter when combined in the majority vote rule gave the highest discrimination accuracy with sensitivity 96.8% 41 and specificity 94.3%. Comparative results for the aggregation rules of min, max, sum, and product are presented in Table 4. 43

## 4. Discussion

In the present study, the proposed system has been designed 45 by combining different classifiers, which were firstly optimized in parameter settings and in the number of features. The panel 47 of classifiers, namely LSMD, Bayesian, k-NN, SVM, and PNN, approached from a different perspective, the problem of sep- 49 arating thyroid nodules into benign and malignant. This was due to the diverse theoretical framework of each algorithm. 51 In this way, complementary information was exploited, which proved valuable for boosting up the system's sensitivity and 53 specificity. An indication supporting this conclusion is that different feature vectors subsets optimized each classifier's per- 55 formance. The Bayesian highlighted as more informative features, those describing nuclei texture. On the other hand, the 57 k-NN, LSMD, PNN, and SVM classifiers optimized their performance using a mixture of morphological and textural fea- 59 tures (see Table 2). Moreover, although the classifiers tested in this study have been successfully applied in computer-assisted 61 microscopy [34–36,43], it has to be stressed that their combination is for the first time investigated. 63

The proposed multi-classifier system was constructed using several combination rules (namely minimum, maximum, aver- 65 age, product, and majority vote) and different mixtures of ensemble classifier members. A comprehensive comparison be- 67



Please cite this article as: A. Daskalakis, et al., Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological..., Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2007.09.005

tween these rules and classifier members combinations was performed in terms of maximizing sensitivity and specificity, using the "test and select" method [32]. The best rule was proved to be the majority vote with ensemble classifier members the PNN, k-NN, and Bayesian. This setup of the multi-classifier system attained significantly higher sensitivity (96.8%) and specificity (94.3%) compared to the best single classifier, which was the PNN in terms of specificity (90.3%) and the SVM in terms of sensitivity (90.3%).

Although statistical approaches [14,44,45] have found significant differences in mean nuclear area and perimeter between groups of patients with benign and malignant thyroid nodules, however, they have concluded that the characterization of individual aspirates based on morphometric features was untrustworthy. On the other hand, more recent studies have explored the potential of pattern recognition methods to overcome these limitations. Karakitsos et al. [8,9] have investigated the use of neural networks and morphometric features in discriminating benign from malignant thyroid lesions, achieving high accuracies of 87% and 98%, employing the specialized May Grunvald Giemsa staining protocol.

In terms of overall classification accuracy, a direct comparison with previous studies is not feasible due to the differences in the data sets and the experimental settings utilized (different staining protocols). Our system was evaluated employing the H&E-staining protocol, which is routinely used in daily clinical practice, although it is not as accurate in labeling nuclei as other specialized staining protocols [8,9]. Thus, in this aspect, our system's performance (95.7% overall accuracy) may be regarded as most encouraging.

Additionally, in terms of clinical use, our system provides a workable framework, since cases were directly classified according to the mean appearance of nuclei extracted from representative areas, indicated interactively by the cytopathologist. Since the main difficulty in diagnostic decision is to verify the existence of malignancy in suspicious areas [3,46], our system was designed to support decision in these areas that may not lead to a definite diagnosis by visual inspection.

## 5. Conclusions

Summarizing, we demonstrated that quantitative analysis of cell nuclei and a suitable combination of different classifiers could improve the performance of single classifiers in diagnosing thyroid nodules. Experimental results are encouraging, illustrating that classifier combination strategies significantly enhance sensitivity and specificity compared to single classifiers. In this way, the proposed system can be utilized in daily clinical practice to support cytopathologists' decisions, when a definite diagnosis is difficult to be obtained. Thus, by using the system as a second opinion tool, excessive reexaminations and unnecessary surgical operations may be avoided.

## Conflict of interest statement

None declared.

## Acknowledgments

The Greek State Scholarships Foundation (I.K.Y.) partly supported this study by funding one of the authors (S.K.).

## References

[1] H. Gharib, Changing concepts in the diagnosis and management of thyroid nodules, Endocrinol. Metab. Clin. North Am. 26 (1997) 777–800.
[2] J.E. Freitas, A.E. Freitas, Thyroid and parathyroid imaging, Semin. Nucl. Med. 24 (1994) 234–245.
[3] O. Ferrer-Roca, E. Ballester-Guardia, J.A. Martin-Rodriguez, Morphometric densitometric and flow cytometric criteria for the automated classification of thyroid lesions, Anal. Quant. Cytol. Histol. 12 (1990) 48–55.
[4] M. Wu, D.E. Burstein, Fine needle aspiration, Cancer Invest. 22 (2004) 620–628.
[5] A. Berner, M. Pradhan, L. Jorgensen, A. Heilo, K.K. Groholt, Fine needle cytology of the thyroid gland, Tidsskr. Nor Laegeforen. 124 (2004) 2359–2361.
[6] E. Artacho-Perula, R. Roldan-Villalobos, F. Blanco-Garcia, A. Blanco-Rodriguez, Objective differential classification of thyroid lesions by nuclear quantitative assessment, Histol. Histopathol. 12 (1997) 425–431.
[7] N. Gupta, C. Sarkar, R. Singh, A.K. Karak, Evaluation of diagnostic efficiency of computerized image analysis based quantitative nuclear parameters in papillary and follicular thyroid tumors using paraffin-embedded tissue sections, Pathol. Oncol. Res. 7 (2001) 46–55.
[8] P. Karakitsos, B. Cochand-Priollet, P.J. Guillausseau, A. Pouliakis, Potential of the back propagation neural network in the morphologic examination of thyroid lesions, Anal. Quant. Cytol. Histol. 18 (1996) 494–500.
[9] P. Karakitsos, B. Cochand-Priollet, A. Pouliakis, P.J. Guillausseau, A. Ioakim-Liossi, Learning vector quantizer in the investigation of thyroid lesions, Anal. Quant. Cytol. Histol. 21 (1999) 201–208.
[10] S. Tseleni-Balafouta, N. Kavantzas, H. Paraskevakou, P. Davaris, Computerized morphometric study on fine needle aspirates of cellular follicular lesions of the thyroid, Anal. Quant. Cytol. Histol. 22 (2000) 323–326.
[11] S. Eldar, E. Sabo, A. Cohen, I. Misselevich, O. Cohen, J. Kelner, C. Mor, J. Shvero, R. Feinmesser, J. Shibi, M. Shabtai, J. Bejar, J.H. Boss, Computer-assisted image analysis of small cell lymphoma of the thyroid gland, Comparison of nuclear parameters of small lymphocytes in lymphomas and Hashimoto's thyroiditis, Comput. Med. Imaging Graph. 22 (1998) 479–488.
[12] T. Nagashima, M. Suzuki, M. Oshida, H. Hashimoto, H. Yagata, T. Shishikura, K. Koda, N. Nakajima, Morphometry in the cytologic evaluation of thyroid follicular lesions, Cancer 84 (1998) 115–118.
[13] H. Galera-Davidson, P.H. Bartels, A. Fernandez-Rodriguez, H.E. Dytch, E. Lerma-Puertas, M. Bibbo, Karyometric marker features in fine needle aspirates of invasive follicular carcinoma of the thyroid, Anal. Quant. Cytol. Histol. 12 (1990) 35–41.
[14] O. Ferrer-Roca, E. Ballester-Guardia, J. Martin, Nuclear chromatin texture to differentiate follicular and papillary carcinoma of the thyroid, Pathol. Res. Pract. 185 (1989) 561–566.
[15] R.A. Ambros, R.C. Trost, A.Y. Campbell, W.C. Lambert, Prognostic value of morphometry in papillary thyroid carcinoma, Hum. Pathol. 20 (1989) 215–218.
[16] G.K. Nguyen, M.W. Lee, J. Ginsberg, T. Wragg, D. Bilodeau, Fine-needle aspiration of the thyroid: an overview, Cytojournal 2 (2005) 12.
[17] Y. Peng, A novel ensemble machine learning for robust microarray data classification, Comput. Biol. Med. 36 (2006) 553–573.
[18] Y. Hayashi, R. Setiono, Combining neural network predictions for medical diagnosis, Comput. Biol. Med. 32 (2002) 237–246.
[19] A. Sboner, C. Eccher, E. Blanzieri, P. Bauer, M. Cristofolini, G. Zumiani, S. Forti, A multiple classifier system for early melanoma diagnosis, Artif. Intell. Med. 27 (2003) 29–44.



Please cite this article as: A. Daskalakis, et al., Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological..., Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2007.09.005

[20] V.L.F. Lum, W.K. Leow, Y. Chen, T.S. Howe, Combining classifiers for bone fracture detection in X-ray images, in: International Conference on Image Processing, 2005.
[21] N. Piliouras, I. Kalatzis, N. Dimitropoulos, D. Cavouras, Development of the cubic least squares mapping linear-kernel support vector machine classifier for improving the characterization of breast lesions on ultrasound, Comput. Med. Imaging Graph. 28 (2004) 247–255.
[22] S. Theodoridis, K. Koutroumbas, Pattern Recognition, Academic Press, Amsterdam, Boston, 2003.
[23] C. Burges, A tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discovery 2 (1998) 121–167.
[24] D. Specht, Probabilistic neural networks, Neural Networks 3 (1980) 109–118.
[25] T. Mouroutis, S. Roberts, A. Bharath, Robust cell nuclei segmentation using statistical modelling, Bioimaging 6 (1998) 79–91.
[26] A. Daskalakis, D. Glotsos, I. Kalatzis, S. Kostopoulos, P. Bougioukos, P. Ravazoula, N. Dimitropoulos, G. Nikiforidis, D. Cavouras, A pixel based classification method for the automatic segmentation of microscopy images, in: First International Conference on Experiments and Process, System Modelling, Simulation and Optimization (1st IC-EpsMsO), Athens, Greece, 2005.
[27] D. Glotsos, P. Spyridonos, D. Cavouras, P. Ravazoula, P.A. Dadioti, G. Nikiforidis, Automated segmentation of routinely hematoxylin–eosin-stained microscopic images by combining support vector machine clustering and active contour models, Anal. Quant. Cytol. Histol. 26 (2004) 331–340.
[28] P. Spyridonos, D. Cavouras, P. Ravazoula, G. Nikiforidis, Neural network-based segmentation and classification system for automated grading of histologic sections of bladder carcinoma, Anal. Quant. Cytol. Histol. 24 (2002) 317–324.
[29] P. Spyridonos, P. Ravazoula, D. Cavouras, K. Berberidis, G. Nikiforidis, Computer-based grading of haematoxylin–eosin stained tissue sections of urinary bladder carcinomas, Med. Inf. Internet Med. 26 (2001) 179–190.
[30] M.M. Galloway, Texture analysis using grey level run lengths, Comput. Graphics Image Process. 4 (1975) 172–179.
[31] R. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man Cybern. 3 (1973) 610–621.
[32] A.J.C. Sharkey, N. Sharkey, U. Gerecke, G.O. Chandroth, The test and select approach to ensemble combination, in: Proceedings of the First International Workshop on Multiple Classifier Systems, 2000.
[33] V.L.F. Lum, W.K. Leow, Y. Chen, T.S. Howe, Combining classifiers for bone fracture detection in X-ray images, in: International Conference on Image Processing, 2005.
[34] C. Decaestecker, I. Salmon, O. Dewitte, I. Camby, P. Van Ham, J.L. Pasteels, J. Brotchi, R. Kiss, Nearest-neighbor classification for identification of aggressive versus nonaggressive low-grade astrocytic tumors by means of image cytometry-generated variables, J. Neurosurg. 86 (1997) 532–537.
[35] L.R. Schad, H.P. Schmitt, C. Oberwittler, W.J. Lorenz, Numerical grading of astrocytomas, Med. Inf. (London) 12 (1987) 11–22.
[36] D. Glotsos, P. Spyridonos, P. Petalas, D. Cavouras, P. Ravazoula, P.A. Dadioti, I. Lekka, G. Nikiforidis, Computer-based malignancy grading of astrocytomas employing a support vector machine classifier, the WHO grading system and the regular hematoxylin-eosin diagnostic staining procedure, Anal. Quant. Cytol. Histol. 26 (2004) 77–83.
[37] J. Kittler, M. Hatef, R.W. Duin, J. Matas, On combining classifiers, IEEE Trans. Pattern Anal. Mach. Intell. 20 (1998) 226–239.
[38] K. Woods, W.P. Kegelmeyer, K. Bowyer, Combination of multiple classifiers using local accuracy estimates, IEEE Trans. Pattern Anal. Mach. Intell. 19 (1997) 405–410.
[39] L.I. Kuncheva, J.C. Bezdek, R.P.W. Duin, Decision templates for multiple classifier fusion: an experimental comparison, Pattern Recognition 34 (2001) 299–314.
[40] L.A. Alexandre, A.C. Campilho, M. Kamel, On combining classifiers using sum and product rules, Pattern Recognition Lett. 22 (2001) 1283–1289.
[41] A. Jain, K. Nandakumar, A. Ross, Score normalization in multimodal biometric systems, Pattern Recognition 38 (2005) 2270–2285.
[42] D. Ruta, B. Gabrys, Classifier selection for majority voting, Inf. Fusion 6 (2005) 63–81.
[43] D. Glotsos, P. Spyridonos, P. Petalas, D. Cavouras, V. Zolota, P. Dadioti, I. Lekka, G. Nikiforidis, A hierarchical decision tree classification scheme for brain tumour astrocytoma grading using support vector machines, in: Proceedings of the Third International Symposium on Image and Signal Processing and Analysis, Rome, Italy, 2003.
[44] Y. Asai, H. Mineta, K. Umemura, A. Yasuhara, H. Mukoudaka, H. Ishizaki, H. Ito, M. Nozue, Statistical analysis on the cases of thyroid papillary carcinoma, Nippon. Jibiinkoka. Gakkai. Kaiho. 96 (1993) 58–65.
[45] K.C. Clark, F.L. Moffat, A.S. Ketcham, A. Legaspi, D.S. Robinson, Nonoperative techniques for tissue diagnosis in the management of thyroid nodules and goiters, Semin. Surg. Oncol. 7 (1991) 76–80.
[46] P. Zagorianakou, V. Malamou-Mitsi, N. Zagorianakou, D. Stefanou, A. Tsatsoulis, N.J. Agnantis, The role of fine-needle aspiration biopsy in the management of patients with thyroid nodules, In Vivo 19 (2005) 605–609.

**Antonis Daskalakis** completed his B.Sc. in Physics from the University of Patras, Greece, in 2003 and his M.Sc. in Medical Physics from the Medical School of the University of Patras in 2005.
Since 2005, he is a Ph.D. candidate in Medical Physics at the University of Patras, Patras, Greece.

**Spiros Kostopoulos** received the B.Sc. degree in Medical Instruments Technology from the Technological Institution of Athens, Athens, Greece, in 2000, and the M.Sc. Degree in Medical Physics from the University of Surrey, Guilford, UK, in 2004.
Since 2005, he is a Ph.D. candidate in Medical Physics at the University of Patras, Patras, Greece, and he is funded by the Greek State Scholarships Foundation.

**Panagiota Spyridonos** holds a B.Sc. degree in Physics, from the University of Patras, Greece. She received her M.Sc. and Ph.D. degrees in Medical Physics from the Medical School of the University of Patras in 2002. She currently holds a post-doc position at the University of Patras in the Department of Medical Physics. Her main research interests include medical image processing, analysis and pattern recognition.

**Dimitris Glotsos** received the B.Sc. degree in Medical Instruments Technology from the Technological Institution of Athens, Athens, Greece, in 2000, and the M.Sc. and Ph.D. degrees in Medical Physics from the University of Patras, Patras, Greece, in 2002 and 2006, respectively.

**Panagiota Ravazoula** is a Medical Doctor working in the Department of Histopathology of the University Hospital of Patras.

**Maria Kardari** is a Medical Doctor working in the Department of Cytology of the University Hospital of Patras.

**Ioannis Kalatzis** completed his B.Sc. in Physics from the University of Athens, Greece, in 1987 and his Ph.D. in Medical Physics from the Medical School of the University of Athens in 2000. During his Ph.D. thesis he was involved in tomographic image processing in the Nuclear Medicine Department of the Hippokrateio General Hospital of Athens, as well in research projects concerning image processing and pattern recognition applications in medicine. He is currently research fellow at the Laboratory of Medical Image and Signal Processing in the Department of Medical Instruments Technology of the Technological Educational Institution of Athens. Dr Kalatzis's research interests include medical image and signal processing and analysis, with focus in the field of pattern recognition.

**Dionisis Cavouras** was born in Kalamata, Greece, in 1951. He received his B.Sc. (1974) in Electronic Engineering, and the M.Sc. (1976) and Ph.D. (1981) in Systems Engineering from the City University, London, UK. He was a research assistant at the Department of Nuclear Medicine, Guy's Hospital, London, UK, between 1976 and 1981; he worked as a Research Fellow at the Department of Computed Tomography, Hellenic Air-force Hospital, Athens, Greece, between 1994 and 1991 and since then he is a professor



Please cite this article as: A. Daskalakis, et al., Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological..., Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2007.09.005

of medical imaging processing at the Department of Medical Instruments Technology in Technological Educational Institution of Athens, and Director of the Laboratory of Medical Image and Signal Processing. He is also an adjunct lecturer on medical image processing at the Department of Informatics, University of Athens, Greece and at the Department of Medical Physics of University of Patras, Greece, and he has been appointed as honorary professor associate at the School of Engineering and Design of Brunel University teaching Network Computing. His research interests include medical image processing, image analysis, pattern recognition, medical statistics, and medical physics. He has published numerous technical and medical papers as well as conference proceedings.

**George C. Nikiforidis** received his Laurea in Physics and his M.Sc. in Atomic and Nuclear Physics both from the University of Milan, Italy, in 1973 and 1980, respectively. He received his Ph.D. in Medical Physics from the University of Patras, Greece, in 1981. He is currently Professor of Medical Physics and Director of the Department of Medical Physics at the University of Patras, where he is also the Dean of the School of Health Sciences and director of the postgraduate course on Medical Physics. He has been the principal investigator or been involved in a variety of national or European research and development projects.



Please cite this article as: A. Daskalakis, et al., Design of a multi-classifier system for discriminating benign from malignant thyroid nodules using routinely H&E-stained cytological..., Comp. Biol. Med. (2007), doi: 10.1016/j.compbiomed.2007.09.005