2nd Hellenic Conf. on AI, SETN-2002, 11-12 April 2002, Thessaloniki, Greece,
Proceedings, Companion Volume, pp. 379-388

Classification of Cell Nuclei Using the Learning Vector Quantizer and Morphological Features

A. Pouliakis[1], P. Karakitsos[2], N. Kalouptsidis[3]

[1] Department of Informatics, Division of Communications and Signal Processing, University of Athens, Panepistimiopolis, T.Y.P.A. Buildings, 15771 ATHENS, GREECE, E-mail: apou@mycosmos.gr.
[2] Department of Histology and Embryology, Medical School of Athens, GREECE. E-mail: pkaraki@cc.uoa.gr
[3] Department of Informatics, Division of Communications and Signal Processing, University of Athens, Panepistimioupolis, T.Y.P.A. Buildings, 15771 Athens, Greece, E-mail: kalou@di.uoa.gr

Abstract. The problem of cell nucleus classification from microscopic images of gastric cytological specimens is considered in this paper. Several variations of the learning vector quantizer (LVQ) are used to perform discrimination into two classes: benign and malignant. Classification is based on morphometric features extracted by a prototype image analysis system (IAS) that was developed specially for cytological laboratories. The presented IAS is used for the image acquisition, segmentation and measurement. The various classifiers are benchmarked and conclusions concerning their behaviour are presented.
The presented feature extraction and classification system can be useful during everyday cytological practice, and also for the treatment of similar cytological problems from various human organs.

KEYWORDS: Neural networks, Learning Vector Quantizer, Image Analysis, Image Segmentation, Image Morphometry, Cytological Diagnosis, Gastric Lesions.

1 Introduction

During the last decade several efforts to evaluate the capability of Artificial Neural Networks (ANNs), in diagnostic cytology have been made [1],[2],[3]. Classification, pattern recognition and decision support are some of the most important and emerging applications of neural nets in the field of cytology. These techniques have not yet attained a wide spread use in cytology for many reasons, the lack of standardization in the specimen preparation and measurements being one of them. Moreover cytologists follow different diagnostic procedures depending on the type of the speciment and the preparation technique. Up to now neural nets and expert systems have been succesfully applied for the mass screening of gynaecological cytology samples [4], for the diagnosis and prognosis of breast lesions [5], urinary lesions [6], gastric lesions [7], and thyroid nodules [3].

The primary cytological diagnostic problem is the discrimination of benign or malignant nature of the nuclei present in the microscopic field of view. The routine procedure includes the following steps: a)sample extraction, b) sample preparation (staining, fixation and in some cases formation of nuclei in a single layer), c) screening, i.e. examination of the sample through microscope and d) decision (i.e. diagnosis).

In the case of gastric samples, cytology has not reached wide acceptance because of the difficulties in the discrimination of benign lesions with severe regenerative alterations from well differentiated cancer cells [8]. Statistical evaluation of morphometric data either at patient level or at the cellular level has not been very successful. The use of neural networks in the above classification problems is very promising. An extension of the system described in this paper could be usefull to the medical society, as in the every day practice of cytological laboratories, are not used similar tools that provide diagnostic consultation. Today cytological laboratories use morphometric and ploidy analysis systems that provide results evaluated by human experts.

2 The Nucleus Measurement System

The complete expert system used during this study appears in fig 1. The image acquisition module acquires microscopic images and transforms them to electrical signals. The image sampling and quantization module translates the electrical signals to digits. The image segmentation subsystem isolates objects (i.e. cell nuclei) on the images. Each object is represented by a vector via the feature extraction subsystem which implements several measurements. Finally each vectorial representation of the nucleus is processed by a group of classifiers that assigns the corresponding nuclei into two categories (benign and malignant). For the specific cytological application the above image processing system was implemented as follows:

Fig. 1. Components of the nucleus measurement imaging system.

The image acquisition subsystem consists of a color CCD camera attached to the top of a microscope. The camera employs a 756×581 CCD sensor. The image sampling and quantization subsystem was implemented by a frame grabber, with 512×512 samples per image and a uniform quantization of 8 bits (i.e. 256 gray levels). The remaining subsystems are implemented by software.

The image segmentation subsystem detects the cell nuclei. The variety of image types make segmentation a difficult task. A mix of manual and completely automated segmentation techniques have been developed, depending on image type and quality of the resulting outcome. Fully automated algorithms are based

on equalization filters and detection of local minima of the image histogram [9]. Automatic segmentation gives satisfactory results if the images do not contain artifacts and the nuclei have similar gray level. Semiautomatic segmentation is based on the histogram's minima at the neighbourhood of the nucleus. When the user points to the center of nucleus, a circular neighborhoud of prespecified radius dictated by everyday practice average nucleus size, is selected. Subsequentlly the region's histogram is extracted, and the minimum is located and used as a threshold to segment the image and extract the selected nucleus. This semiautomatic technique is suitable for nuclei that have relatively uniform distribution of gray level. If the previous algorithms does not give good results, the user has to draw the nucleus boundary by a pointing device such as a mouse. Factors that affect the image segmentation are the specimen staining procedure that affecs the color and thus the gray values of nuclei, the clumping of cells and some diseases that cause alteration of the nucleus characteristics making the detection difficult.

A variety of features are extracted by the system. These are based on factors that affect the decision during screening: a) the size of cell nucleus b) the shape c) the texture d) the color e) the homogeneity of nuclei characteristics at the specimen and f) contextual characteristics. For the presented study, the data set can be downloaded from the following URL: http://www.di.uoa.gr/~makis/projects/CCS/CCS.html.

Table 1 summarizes the features generated by the extraction system for each nucleus based on standard cytological practice. Features are grouped according to nucleus characteristics into *geometric* and *densitometric*. Densitometric features are mainly associated with texture.

- The histogram $h(k), k = 1...N$ shows the number of pixels with intensity k, inside the Region Of Interest (ROI).
- The optical density $x(i,j)$ denotes the gray level of pixel (i,j), L defines the ROI, and N is the number of pixels in the ROI. Note that the opical density is equal to the mean value of the histogram
- The histogram of diferences: $\mathbf{d} = (d_1, d_2)$ is a displacement vector in the $2D$ space, $diff = |x(i,j) - x(i-d_1, j-d_2)|$ is the diference between the pixel at (i,j) with value $x(i,j)$ and the pixel at distance $\mathbf{d}$ from (i,j). The histogram of diferences $h_d(k), k = 1...N$ is calculated in two steps: 1) calculation of the diference for all pixels that belong to the object for the specific displacement vector and 2) calculation of the histogram of the differences found in step 1.
- Co-occurrence matrix: $C(i,j)$ with dimensions $N \times N$.
- Run Length Matrix: $R(i,j), M \times K$ with $K = \sqrt{L^2 + M^2}$ for images of dimensions $L \times N$ with M gray levels.

The geometric features are calculated via the boundary of a nucleus. They describe properties relevant to size (for example area, perimeter, diameter) and shape (eg. FormAR, FormPE, circularity). A detailed description of the computational methods employed to determine geometric characteristics is supplied in [10] and [11].

Table 1. Nucleus features measured by the IAS.

Geometric features	
Area	A
Perimeter	P
Major axis	Mj
Minor axis	Mn
Diameter	$D = \sqrt{\frac{2A}{\pi}}$
Circularity	$Curl = \frac{(Mj)-(Mn)}{(Mj)}$
Roundness factor	$RF = \frac{P}{2\sqrt{\pi(A)}}$
Contour ratio	$CR = \frac{(P)^2}{4\pi(A)}$
Contour index	$CI = (P)\sqrt{A}$
Form area (FormAR)	$FormAR = \frac{1}{4\pi(Mj)(Mn)}$
Form perimeter (FormPE)	$FormPE = 4\pi(A)(P)^2$
Mean Radius of nucleus	
Textural features	
Optical density	$OD = \frac{\sum\sum_{(i,j)\in L} x(i,j)}{N}$, $\mu = \frac{\sum_{k=1}^{N} kh(k)}{N}$
Standard deviation of histogram	$STD(h(k))$
Variance of histogram	$VAR(h(k)) = \frac{\sum_{k=1}^{N}(h(k)-\mu)^2 h(k)}{N^2}$
Short run of run length matrix	$\frac{\sum_{i=1}^{M}\sum_{j=1}^{K}\left(R(i,j)/K^2\right)}{\sum_{i=1}^{M}\sum_{j=1}^{K}(R(i,j))}$
Long run of run length matrix	$\frac{\sum_{i=1}^{M}\sum_{j=1}^{K}\left(K^2 R(i,j)\right)}{\sum_{i=1}^{M}\sum_{j=1}^{K}(R(i,j))}$
Grey level of run length matrix	$\frac{\sum_{i=1}^{M}\left(\sum_{j=1}^{K} R(i,j)\right)^2}{\sum_{i=1}^{M}\sum_{j=1}^{K}(R(i,j))}$
Distribution of run length matrix	$\frac{\sum_{j=1}^{K}\left(\sum_{i=1}^{M} R(i,j)\right)^2}{\sum_{i=1}^{M}\sum_{j=1}^{K}(R(i,j))}$
Maximum of co-occurrence matrix	$Max(C(i,j))$
Inertia of co-occurrence matrix	$I_d = \sum_{k=1}^{N}\sum_{l=1}^{N} \lvert k-l\rvert^2 C(k,l)$
Entropy of co-occurrence matrix	$C_d = \sum_{k=1}^{N}\sum_{l=1}^{N} C(k,l)\ln(C(k,l))$
Contrast of differences histogram	
Mean value of differences histogram	$\mu_d = \frac{\sum_{k=1}^{N} kdiff(k)}{N}$
Standard deviation of differences histogram	$STD(h_d(k))$
Entropy of differences histogram	$Entr(h_d(k)) = \frac{\sum_{k=1}^{N} h_d(k) ln(h_d(k))}{-N^3}$

Densitometric features are relevant to the values of the pixels inside the nucleus region and are relevant to texture caused by the nucleus chromatin. From the various methods proposed in the literature for textural descriptors (see [9] and [12]), four models have been applied, based on: a) histogram b) differences histogram c) run length matrix and d) co-occurrence matrix. The first two models are computationally simple but texture discrimination is poor. The other two models are more compex but give better information about the texture structure [9]. A detailed description of the computational methods of textural features is provided by Pitas in [9].

3 Classification and data set description

Bipartite classification into benign and malignant nuclei is considered. Three variants of the LVQ algorithm [13],[14] were evaluated. Kohonen et. al. [15] have developed a program package to implement the three variations [16] of the LVQ algorithm (LVQ1, LVQ2.1, and Optimized LVQ1 "OLVQ1") .

The data set (see table 2) includes 120 patient cases. The number of vectors for each case ranges 35 to 183 with a mean value of 120. Five pathological categories are studied: ulcer, gastritis, inflamatory displasia, cancer and true displasia. The data set is divided into the training set (30% of the data set) and the test set. The corresponding distribution of the training set reflects the distribution of the test set (see table 2). The assignment of a vector into the training or test set is random. However 30% of vectors from each case are used in the training set.

To assess the robustness of the algorithms with respect to the training set selection, ten different training and test sets were used by random assignment.

The displasia types constitute the most difficult cases for cytological diagnosis. In practice these are diagnosed after further examinations.

The classification schemes assign vectors into two classes: benign (cancer and true displasia) and malignant (gastritis, ulcer and inflamatory displasia).

Table 2. Data set description

Class	Cases	Measured nuclei	Training vectors	Test vectors
cancer	25	2920	876	2044
true displasia	5	370	111	259
gastritis	25	3150	945	2205
ulcer	60	6550	1965	4585
inflamatory displasia	5	310	93	257
total	120	13300	3990	9310

3.1 Training of the LVQ networks

The main parameter for configuring the LVQ arcitecture is the number of codevectors. This is determined by simulation and computer search. A steady linear increase was noticed once the number of codevectors exceeds 500. Note that for the test set, the class +1 (malignant) accuracy is considerably less than the accuracy for class −1 (benign). LVQ1 and LVQ2.1 algorithms exhibit similar behavior.

In the sequel we explored the behaviour of the classifiers on the training and test set when using different codevectors. We notice almost linear increase on the test set when the accuracy increases on the training set. Similar behaviour is noticed also for the LVQ1 and LVQ2 algorithms, a fact indicating that there was no over-training. This result justifies the selection of the network that will be finally employed, as the one that has the better accuracy on the training set, as it is expected that the same network will give better accuracy on the test set.

The comparison of the performance of the three algorithms indicates that OLVQ1 provides superior overall accuracy both for the training and the test set.

In a previous work of the same authors [7], there were obtained results using the back propagation algorithm on a subset of this dataset. The results of that method were similar to the ones obtained here (95.7% − 97.3% accuracy on the test set). However training of neural networks using the back propagation algorithm requires exhausting experiments to obtain the best classifier and fine tune it's parameters, training of each classifier requires much more computational power than each LVQ type architecture.

Finally in order to test the robustness of each one of the LVQ1, LVQ2.1 and OLVQ1 algorithms we construct the confusion matrices based on the ten different training and test sets. The confusion matrices for the training and test sets for the three algorithms appear in tables 3, 4 and 5.

Table 3. Confusion matrices of LVQ1 classifier for the training and test set. Mean Overall Accuracy $94.97 \pm 0.29\%$ and $93.03 \pm 0.22\%$ respectivelly

	Training set		**Test set**	
	Benign	**Malignant**	**Benign**	**Malignant**
Benign	95.20 ± 0.34	4.80 ± 0.34	93.88 ± 0.40	6.12 ± 0.40
Malignant	5.72 ± 0.22	94.28 ± 0.22	9.58 ± 0.82	90.42 ± 0.82

Table 4. Confusion matrices of LVQ2.1 classifier for the training and test set. Mean Overall Accuracy $94.91 \pm 0.30\%$ and $93.01 \pm 0.22\%$ respectivelly.

	Training set		**Test set**	
	Benign	**Malignant**	**Benign**	**Malignant**
Benign	95.14 ± 0.32	4.86 ± 0.32	93.85 ± 0.41	6.15 ± 0.41
Malignant	5.78 ± 0.32	94.22 ± 0.28	9.56 ± 0.80	90.44 ± 0.80

Table 5. Confusion matrices of OLVQ1 classifier for the training and test set. Mean Overall Accuracy $95.36 \pm 0.42\%$ and $93.15 \pm 0.19\%$ respectivelly.

	Training set		**Test set**	
	Benign	**Malignant**	**Benign**	**Malignant**
Benign	95.68 ± 0.44	4.32 ± 0.44	94.09 ± 0.38	5.91 ± 0.38
Malignant	5.62 ± 0.44	94.38 ± 0.49	9.71 ± 0.87	90.29 ± 0.87

4 Conclusions

The main conclusions derived are: The majority of misclassified nuclei (about 90%) are extracted from specific patient cases (17.5% of the cases). This fact may be attributed to (a) inherent case misclassification in the human classification phase by the cytologists or (b) the majority of the nuclei measured for these cases have peculiar characteristics that forces assignment to the wrong class. Figure 2 shows several misclassified nuclei, their appearance concerning the geometric characteristics and especially the textural characteristics are quite similar, a fact that confirms that there is an overlap in the feature space. Moreover most nuclei are from displasia cases. It is worth noting that two cases in the data set were given as uncorfimed true displasias, but finally and after these results, these cases were assigned as inflammatory displasias. By further analysis there were discovered 11 cases that have numerous misclassified nuclei. Out of these cases, 4 are displasias (i.e. 36%). However the displasias in the complete data set are 10 in 120 (i.e. 8%). This indicates that the open cytological problem of classifying displasias, remains for the other classification methods. This fact does not limit the merits of the proposed method, as the obtained overall accuracy is significantly better than standard cytological screening. Moreover it does not require an additional examination to clarify displasias.

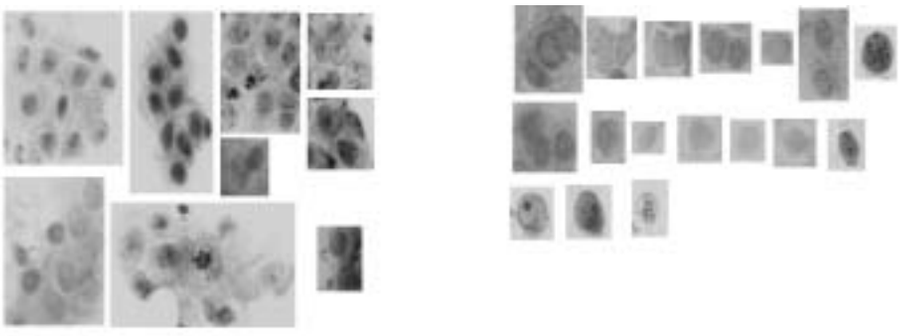

Fig. 2. Misclassified nuclei. Left: false positive, right: false negative.

By the various confusion matrices produced, and by the training/testing procedure many observations may be done:

- LVQ classifiers yield better accuracy on the training set than on the test set.
- The LVQ classifiers have better accuracy on data extracted from malignant nuclei than back propagation [7].
- Training of the LVQ classifiers, requires seconds or minutes, in contrast training of other algorithms such as the back propagation requires more training time.
- Robustness tests of classification accuracy, indicates that LVQ classifiers yield very low standard deviation.

The statistics of the best classifier for nucleus classification are as follows: Specificity=94.23%, Sensitivity=89.51%, Predicted Value of Positive Result=83.59%, Predicted Value of Negative Result=96.47%, Overall Accuracy=93.06%. The calculations are based on classification results for the entire data set.

5 Discussion

Cytology has not reached wide acceptance in the investigation of gastric lesions because of the high rate of false positive and negative results. Neural networks as shown in this paper have the potential for changing this attitude. The generalization at the individual patient level either via a majority logic technique or by directly applying NNs to statistical quantities of measures obtained by nuclei from a single patient case, are both very promissing for classification purposes.

In summary: The use of morphometry to measure size, shape and texture characteristics of cell nucleus during microscopic screening, enables us to quantify the characteristics that during routine laboratory work are subjectively examined. NNs give excellent results in most cases at the nuclear level. Integration of conventional cytological screening with measurement-classification processes seem to increase the quality of today cytological laboratory.

Extensions if this paper should focus on exhausting tests and validation on very large data sets so as to assure the robustness of the method, usage of contextual features (e.g. nuclei density within a predefined area of the specimen), and improvement of the segmentation methods towards an automated nuclei detection system.

References

1. Astion, M.L., Wilding, P.: The application of back propagation neural networks to problems in pathology and laboratory medicine. Arch. Pathol. Lab. Med. **116** (1992) 995–1001
2. Astion, M.L., Wilding, P.: Application of neural networks to the interpretation of laboratory data in cancer diagnosis. Clin. Chem. **38** (1992) 34–38
3. Karakitsos, P., Cochand-Priollet, B., Guillausseau, P., Pouliakis, A.: Potential of the back propagation neural network in the morphologic examination of thyroid lesions. Analytical and Quantitative Cytology and Histology **18** (1996) 494–500

4. Karakitsos, P., Voulgaris, Z., Pouliakis, A., Kyroudes, A., Stergiou, E.B., Ioakim-Liossi, A., Kittas, C.: The potential of the learning vector quantizer in the cell classification of endometrial lesions from post menopausal women. 11th International meeting on gynaecological Oncology (1999) Budapest, Hungary.
5. Wolberg, W.H., Mangassarian, O.L.: Computer-aided diagnosis of breast aspirates via expert systems. Anal. Quant. Cytol. Histol. **12** (1990) 314–320
6. Pantazopoulos, D., Karakitsos, P., Iokim-Liossi, A., Pouliakis, A., Dimopoulos, K.: Comparing neural networks in the discrimination of benign from malignant low urinary lesions. British Journal of Urology **81** (1998) 574–579
7. Karakitsos, P., Stergiou, E.B., Pouliakis, A., Tzivras, M., Archimandritis, A., Liossi, A.I., Kyrkou, K.: The potential of the back propagation neural network in the discrimination of benign from malignant gastric cells. Analytical and Quantitative Cytology and Histology **18** (1996) 245–250
8. Kasugai, T., Kobayashi, S.: Evaluation of biopsy and cytology in the diagnosis of gastric cancer. Am. J. Gastrenterol **62** (1974) 199–203
9. Pitas, I.: Digital Image Processing Algorithms. First edn. Prentice-Hall (1993)
10. Introduction to cytometry and histometry. Technical report, 1st COMETT International Cource on Microscope Imaging in Biology and Medicine, Grenoble, France (1991)
11. Russ, J.C.: The Image Processing Handbook. Second edn. CRC Press, IEEE Press, Boka Raton (1995)
12. Sonka, M., Hlavac, V., Boyle, R.: Image processing analysis and machine vision. First edn. Chapman&Hall, Cambridge, Great Britain (1994)
13. Kalouptsidis, N.: Signal Processing Systems, Theory and Design. John Wiley & Sons Inc. (1997)
14. Kohonen, T.: Self-Organization and Associative Memory. Third edn. Springer-Verlag, New York (1988)
15. Kohonen, T., Kangas, J., Laaksonen, J., Torkkola, K.: LVQ PAK: The learning vector quantization program package, version 2.1, Rakentajanaukio, Finland. (1993)
16. Kohonen, T.: Improved versions of learning vector quantization. International Joint Conference on Neural Networks, San Diego CA. (**1**) 545–550