Applying computational intelligence methods for tumour characterization

E.I. Papageorgiou[1], G. Georgoulas[1], C.D. Stylios[2] and P.P. Groumpos[1]

[1] Department of Electrical & Computer Engineering, University of Patras, Rion Patras, Greece
[2] Department of Communications, Informatics & Management, TEI of Epirus, Artas Epirus, Greece

Abstract— **For most tumor types in histopathology, including bladder tumors, the morphology of tissues-as viewed through a light microscope-express the tumor characterization as malignant or not. Due to the subjective nature of this classification, a large number of computer assisted methodologies have been used to increase the diagnostic and/or prognostic value of tumors' characterization. In this research work, computational intelligence methods such as Fuzzy Cognitive Maps (FCMs) and Support Vector Machines (SVMs) are explored for assisting tumour characterization. The classification results are presented and are compared with other conventional classifiers proving their efficiency.**

Keywords— **Decision making, Fuzzy cognitive maps, Support vector machine, statistical learning, bladder tumor grading, classification**

I. Introduction

Decision analysis consists of a number of quantitative methods to aid in choosing among alternative decisions [1]. Traditional decision analysis is used to indicate decisions favouring good outcomes even though there is risk surrounding the decision. Furthermore, the value of each possible outcome of a decision, whether measured in costs and benefits or utility, is usually variable. The traditional decision analysis, in combination with sensitivity analysis, has become a standard methodology for using existing data and expert opinion to examine effectiveness and cost-effectiveness issues in health care [2].

Over the last years, several approaches have been proposed attempting to hybridise machine learning techniques with knowledge based methods. Fuzzy Cognitive Map (FCM) is an advanced methodology used for modeling knowledge based systems providing flexibility on the simulated system's design and control. This flexibility of FCMs, as well as their learning properties, make their choice attractive for a variety of modeling and support decision tasks.

Support Vector Machines pioneered by Vapnik [3] are universal feedforward networks that can be used both for pattern classification and nonilinear regression. In this work are combined with FCMs to achieve tumor malignancy classification.

In a previous work, a Fuzzy Cognitive Map model for tumor grading, namely Fuzzy Cognitive Map Grading Tool (FCM-GT), has been proposed consisting on two stage processes: on the first stage accumulate, model and process expert knowledge and on the second stage, characterize tumors based on quantitative data [4]. The tumor grading was based on the histopathological characteristics, which were determined as the concepts of the FCM model which in turn was trained using an unsupervised non-linear Hebbian learning algorithm [5,6].

This work presents the classification results for both computational intelligence techniques which seem very promising and compares them with other conventional classifiers.

II. Support Vector Machines and Fuzzy Cognitive Maps

A. Description of Support Vector Machines

Support Vector Machines are in fact linear machines which are trained using an algorithm based on optimization theory [3], [7-12]. For a two class problem, an SVM is trained so that the decision function maximizes the generalization ability. For a separable classification task the idea is to map the training data from the m-dimensional input space into an n-dimensional $(n > m)$ feature space using a kernel function *K,* where a separating hyperplane hyperplane can be found by solving a quadratic programming problem.

In case of non-separable problems, as it is the case in real world problems, the SVM solution finds the hyperplane in feature space that keeps both the empirical error small and maximizes the margin between the hyperplane and the instances closest to it. This can be done by minimizing:

$$\frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i \tag{1}$$

subject to

$$y_i\left(\left(\mathbf{w}\cdot\phi(\mathbf{x}_i)\right)+b\right)\geq 1-\xi_i \quad \xi_i \geq 0 \quad i=1,\ldots,n \tag{2}$$

where ξ_i are slack variables, which are introduced to allow the margin constraints to be violated, and $\boldsymbol{\varphi}(\cdot)$ is the nonlinear mapping from the input space to the feature space. Parameter C controls the trade off between maximizing the

margin and minimizing the error and it is usually determined through a cross-validation scheme [8-10].

The class prediction for an instance $\mathbf{x}$ is given by:

$$f(\mathbf{x}) = sign\left(\sum_{i=1}^{l} y_i \alpha_i K(\mathbf{x}_i, \mathbf{x}) + b\right) \quad (3)$$

where the coefficients a_i are calculated by maximizing the Lagrangian:

$$\sum_{i=1}^{l} a_i - \frac{1}{2}\sum_{i,j=1}^{l} a_i a_j y_i y_j K(\mathbf{x}_i, \mathbf{x}) \quad (4)$$

subject to $\sum_{i=1}^{l} y_i \alpha_i = 0$ and $0 \le \alpha_i \le C$, $i = 1, 2, ..., l$

and K is called the inner-product kernel [7-12].

One of the most commonly used kernel is the Radial Basis Function (RBF) kernel:

$$K(\mathbf{x}, \mathbf{x}_i) = \exp\left(-\frac{1}{2\sigma^2}\|\mathbf{x} - \mathbf{x}_i\|^2\right) \quad (6)$$

It must be mentioned that even though, SVMs were primarily designed for two-class, classification problems. However there are methods there have been proposed methods in order to deal with multi-class classification [12]. In this work we adopted the pairwise classification (or one to one) approach.

B. Brief Description of Fuzzy Cognitive Maps

FCM is a soft computing technique that follows an approach similar to the human reasoning and human decision-making process. FCM consists of nodes (concepts) that illustrate the different aspects of the system's behavior. These nodes (concepts) interact with each other showing the dynamics of the model. Generally, concepts reflect attributes, characteristics, qualities and senses of the system. Each concept represents one of the key-factors of the modeled system and its value is represented by a number A_i. Interconnections among concepts of FCM signify the cause and effect relationship that a concept has on the others [13].

FCM is developed by human experts who operate/supervise/know the system and its behavior under different circumstances in such a way that the accumulated experience and knowledge are integrated in a causal relationship between factors/characteristics of the system.

The structure of the FCM model can be viewed as a recurrent artificial neural network, where concepts are represented by neurons and causal relationships between concepts by weighted links connecting the neurons. These weighted interconnections represent the direction and degree with which a concept influences the value of the interconnected concepts. Each interconnection between two concepts C_i and C_j, has a weight w_{ij}, belonging to the interval [-1, 1]. The sign of w_{ij} indicates whether the relation between the two concepts is direct or inverse. The methodology for developing FCMs is given in [14] and it is based on a group of experts who are asked to define concepts and describe relationships among concepts.

In general, the values of each concept are calculated by collecting-aggregating the influence of the other concepts to the specific one, by applying the following rule:

$$A_i(k+1) = f(A_i(k) + \sum_j A_j \cdot w_{ji}) \quad (7)$$

where $A_i(k+1)$ is the value of concept C_i at time k+1, $A_j(k)$ is the value of concept C_j at time k, and f is a sigmoid threshold function.

To increase the robustness and efficiency of the FCM model, learning techniques have been proposed. These learning methods modify the initial FCM weights according to the system inputs and specifications [6,15]. Unsupervised learning algorithms have been used effectively giving reliable results in the task of FCM tumor grading [4].

C. Description of Fuzzy Cognitive Map Grading Tool

The experts who asked to develop and construct the FCM model for tumor grading using the methodology presented previously [4], were histopathologists with deep knowledge and clinical experience. The doctors-histopathologists defined the main histopathological features (concepts) that play important role in the final grade characterization. More specifically, eight well documented in the bibliography histopathological criteria (features) essential for tumour grading (Table1) [16-19] were used and each tissue section (patient slide) was evaluated retrospectively by histopathologists using these features. These considered features are the causative variables or factors of the tumour grading system that have been selected by experts to construct the FCM for tumour grading [4], [20].

The FCM tumour grading model was developed consisting of the following 9 concepts: Concept C_1 represents the cell distribution, C_2 represents the cell size, C_3 the cell number, C_4 the cytoplasm, C_5 the nuclei, C_6 the nucleoli, C_7 the necrosis, C_8 the mitoses and C_9 the degree of tumour grade.

Table 1 Main factors for tumor grading

Histological feature	Possible Assessment
Cell distribution	Even, clustered
Cell size	Uniform, pleomorphic
Cell number	Numerous, variable
Cytoplasm	Homogeneous, variable
Nuclei	Uniform, irregular, very irregular, bizarre
Nucleoli	Inconspicuous, evident, prominent
Necrosis	Inconspicuous, frequent
Mitosis	Absent-rate, occasional, numerous

In order to use the FCM model histopathologists were asked to examine each tissue section retrospectively and estimate the value of the eight histopathological variables (Table 1); these values were transformed in the range [0, 1], and were assigned to the corresponding concepts [20].

Then, for weight matrix determination, histopathologists were described the cause-effect relationships among these concepts using IF-THEN rules and through defuzzification method of GoA, the initial weights of the FCM tumor-grading model were determined (see Figure 1).

The tumor grading procedure is mainly oriented towards the determination of the values of two concepts: concept "Nuclei" which represents the characteristics of nuclear appearance and may be considered that figures out the nuclear grading concept, and concept "Grade" that characterizes the degree of tumor malignancy. These two concepts are considered the main attributes for the final degree of tumor characterization.

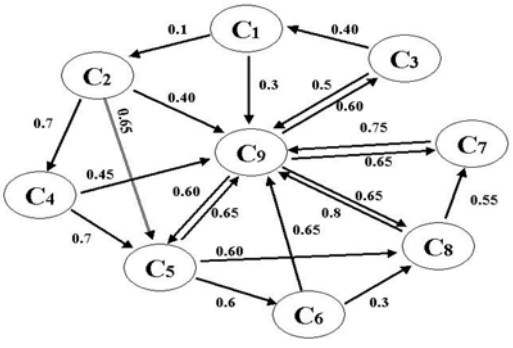

Fig. 1 The FCM tumor grading model

The unsupervised NHL algorithm is used to modify the weights of the FCM grading model for each examined case of urinary bladder tumors. According to the NHL algorithm, experts were asked to select the input and output concepts that determine the triggering process. The concepts C_5 and C_9 were defined as the Decision Output Concepts (DOCs), which determine the main aspects of tumor grade. All the other concepts have been defined as input concepts.

III. Applying Support Vector Machines for Bladder Tumor Grading

The SVMs were combined with the FCM tumor-grading model to assist bladder tumor diagnosis. After the development of the FCM model and the determination of the necessary specifications for the implementation of the NHL algorithm, the training procedure was applied and then, through the use of a SVM with RBF kernels the grade of the tumor was assigned.

The experimental data consisted of one hundred twenty-eight tissue sections (slides) for urinary bladder with superficial transitional cell carcinoma were retrieved from the archives of the Department of Pathology of University Hospital of Patras, Greece. Tissue sections were routinely stained with Haematoxylin-Eosin. Every case was reviewed independently by the doctors-experts to safeguard reproducibility. Histopathologists had classified the cases following the World Health Organization (WHO) grading system as follows: forty cases as grade I, forty-three as grade II, and forty-five as grade III.

Since there was significant variance between experts in recognising and evaluating certain histopathological patterns in a specimen, only one expert was asked to examine each tissue section and estimate the values of the eight histopathological variables (Table 1). Due to the restricted number of cases, we used the leave one out method [20] in order to estimate the performance of the classification method, using 127 of the cases to build the model and testing it on the example left out. Moreover in order to avoid using the same data set for both building the model and estimating its performance [21,22], for each one of the 128 subsets we employed again the leave one out method within the 127 cases that consisted the training set in order to select the model's parameters. Once the parameters were optimized the SVMs was retrained for the whole training set (127 cases) and its performance was evaluated using the corresponding case that was originally left out. For comparison reasons we also tested the performance of the nearest neighbor [23] and the results are summarized in Table 1.

In a previous work, our research group has introduced the FCM-grading tool, working exclusively on the qualitative assessments of histopathological variables [4]. This work extends this research line exploring means to improve classification accuracy by utilizing SVM classifier to assist medical diagnosis.

Table 1 Comparative results

Tumor category	FCM-GT	NNs	FCM-SVMs	k-nn	PROAFTN
Grade I	72,55%	64,9%	80%	87.5%	71.13%
Grade II	74,42%	69,4%	65.1%	58.1%	52%
Grade III	95,55%	82,7%	77.8%	71.1%	57.73%
Overall Accuracy			74.22	71.9%	

IV. Conclusions

In this research effort, computational intelligence methods were examined for characterizing bladder tumors and compared with other convenient classifiers proving efficient results. The SVM classifier was introduced in the FCM tumor grading model to enhance the characterization of urinary bladder tumors. The advanced computational technique of FCM model with SVMs exhibited high performance in correctly classifying Grade I tumors outperforming FCM-GT, NNs and PROAFTN [24]. However its performance concerning Grade II and Grade III tumors is comparable to NNs and worse compared to FCM-GT even though it is better compared to PROAFTN. Therefore, although the proposed method improves the performance compared to PROAFTN it has to be further developed and investigated since hasn't managed to give better results compared to FCM-GT except from the case of the Grade I tumors.

Acknowledgment

The work of E.I. Papageorgiou was supported by a post-doctoral research grant from the Greek State Scholarship Foundation (I.K.Y.).

References

1. Sox J H C, Blatt M A, Higgins M C, Marton K I (1988) Medical Decision Making, Boston, Massachusetts: Butterworths.
2. Dietterich T G (1998). Machine-learning research: Four current directions. The AI Magazine 18(4):97–136 URL: http://citeseer.nj.nec.com/dietterich97machine.html
3. Vapnik V N (1995) The nature of statistical learning theory. New York: Springer-Verlag.
4. Papageorgiou E, Spyridonos P, Ravazoula P, Stylios C D, Groumpos P, Nikiforidis G (2006) Advanced Soft Computing Diagnosis Method for Tumor Grading, Artif Intell Med, 36:59-70.
5. Papageorgiou E, Groumpos P (2005) A weight adaptation method for fine-tuning Fuzzy Cognitive Map causal links, Soft Computing Journal, 9: 846-857.
6. Papageorgiou E I, Spyridonos P, Stylios C D, Nikiforidis G, Groumpos P P (2004) The Challenge of Using Soft Computing Techniques for Tumor Characterization, *16th Intern. Conf. on Artif Intel & Soft Comp,* Poland, 7-11 June, LNCS 3070:1031-1036.
7. Burges C J C (1998) A Tutorial on Support Vector Machines for Pattern Recognition, Data Mining and Knowledge Discovery, 2:121-167.
8. Muller, K. R., Mika, S., Ratsch, G., et al. (2001) An Introduction to Kernel-Based Learning Algorithms. IEEE Trans. Neural Nets, 2:181-201.
9. Veropoulos K, Cambell C, Cristianini N (1999) Controlling the sensitivity of support machines, Proc. Int. Joint Conf. on Artificial Intelligence (IJCAI99), Stockholm, 55-60.
10. Cristianini N, Shawe-Taylor J (2000) An Introduction to Support Vector Machines. Cambridge University Press,.
11. Osuna E E, Freund R, Girosi F (1997) Support Vector Machines: Training and Applications, MIT, A.I. Memo No. 1602.
12. Scholkopf B, Smola A J (2002) Learning with Kernels. Support Vector Machines, Regularization, Optimization, and Beyond. MIT Press, England
13. Kosko B (1992) Neural Networks and Fuzzy Systems, Prentice-Hall, New Jersey.
14. Stylios C, Groumpos P (2004) Modeling Complex Systems using Fuzzy Cognitive Maps. IEEE Trans Syst Man Cyber 34:159-165
15. Papageorgiou E I, Stylios C D, Groumpos P (2004) Active Hebbian Learning to Train Fuzzy Cognitive Maps. Int. J. Approx. Reasoning, 37:219-249.
16. Parker S L, Tony T, Bolden S, Wingo P A (1997) Cancer Statistics. CA Cancer J Clinics, 47 (5):5-27.
17. Ooms E, Anderson W, Alons C, Boon M, Veldhuizen R (1983) Analysis of the performance of pathologists in grading of bladder tumours. Human Pathology, 14:140-143.
18. Bostwick G, Ramnani D, Cheng L (1999) Diagnosis and grading of bladder cancer and associated lesions. Urologic Clinics of North America. 26:493-507.
19. Gil J, Wu H, Wang B Y (2002) Image analysis and morphometry in the diagnosis of breast cancer. Microsc. Res. Techniq. 59:109-118.
20. Spyridonos P, Ravazoula P, Cavouras D, Berberidis K, Nikiforidis G. (2001) Computer-based grading of haematoxylin-eosin stained tissue sections of urinary bladder carcinomas, Med Inform & Intern Med, 26 3:179-190.
21. Duda R O, Hart P E, and Stork D G (2001) Pattern Classification. 2nd ed. John Wiley & Sons
22. Salzberg S L (1997) On Comparing Classifiers: Pitfalls to avoid and a Recommended Approach, Data Mining and Knowledge Discovery, 1: 317-328.
23. Michie D, Spiegelhalter D J, Taylor C C (1994) Machine Learning, Neural and Statistical Classification. Englewood Cliffs, N.J.: Prentice Hall. Data available at: http://www.ncc.up.pt/liacc/ML/statlog/datasets.html.
24. Belacel N, Boulassel M (2001) Multicriteria fuzzy assignment method: a useful tool to assist medical diagnosis. Artif Intell Med 21:201-207

Address of the corresponding author:

Author: Elpiniki Papageorgiou
Institute: University of Patras
Street: Eratosthenous 6
City: Patras
Country: Greece
Email: epapageo@ee.upatras.gr