Artificial Intelligence in Medicine (2006) **36**, 59–70

ELSEVIER

ARTIFICIAL INTELLIGENCE IN MEDICINE

http://www.intl.elsevierhealth.com/journals/aiim

Advanced soft computing diagnosis method for tumour grading

E.I. Papageorgiou [a,*], P.P. Spyridonos [c], C.D. Stylios [b,1], P. Ravazoula [d], P.P. Groumpos [a], G.N. Nikiforidis [c]

[a] *Department of Electrical and Computer Engineering, Laboratory for Automation and Robotic, University of Patras, Rion 26500, Greece*
[b] *Department of Communications, Informatics and Management, TEI of Epirus, 47100 Kostakioi, Artas, Epirus, Greece*
[c] *Computer Laboratory, School of Medicine, University of Patras, Rion 26500, Greece*
[d] *Department of Pathology, University Hospital of Patras, Rion 26500, Greece*

Received 20 November 2003; received in revised form 9 March 2005; accepted 22 April 2005

KEYWORDS
Soft computing;
Fuzzy cognitive maps;
Grading urinary bladder tumours;
Active Hebbian learning algorithm

Summary

Objective: To develop an advanced diagnostic method for urinary bladder tumour grading. A novel soft computing modelling methodology based on the augmentation of fuzzy cognitive maps (FCMs) with the unsupervised active Hebbian learning (AHL) algorithm is applied.
Material and methods: One hundred and twenty-eight cases of urinary bladder cancer were retrieved from the archives of the Department of Histopathology, University Hospital of Patras, Greece. All tumours had been characterized according to the classical World Health Organization (WHO) grading system. To design the FCM model for tumour grading, three experts histopathologists defined the main histopathological features (concepts) and their impact on grade characterization. The resulted FCM model consisted of nine concepts. Eight concepts represented the main histopathological features for tumour grading. The ninth concept represented the tumour grade. To increase the classification ability of the FCM model, the AHL algorithm was applied to adjust the weights of the FCM.
Results: The proposed FCM grading model achieved a classification accuracy of 72.5%, 74.42% and 95.55% for tumours of grades I, II and III, respectively.
Conclusions: An advanced computerized method to support tumour grade diagnosis decision was proposed and developed. The novelty of the method is based on employing the soft computing method of FCMs to represent specialized knowledge

* Corresponding author. Tel.: +30 2 610 997293; fax: +30 2 610 997309.
E-mail addresses: epapageo@ee.upatras.gr, groumpos@ee.upatras.gr (E.I. Papageorgiou), stylios@teleinfom.teiep.gr (C.D. Stylios).
[1] Tel.: +30 2681026825; fax: +30 2681075839.

0933-3657/$ – see front matter
doi:10.1016/j.artmed.2005.04.001

on histopathology and on augmenting FCMs ability using an unsupervised learning algorithm, the AHL. The proposed method performs with reasonably high accuracy compared to other existing methods and at the same time meets the physicians' requirements for transparency and explicability.

1. Introduction

Traditionally, histological examination for the classification of tumours has been based on the morphology of tissues inspected through a light microscope. Classification has great significance because the modality of therapy for urinary bladder tumours highly depends on the morphological tumour characterization [1]. In 1973, the World Health Organization (WHO) introduced standards according to which tumours are classified as grade I, II or III [2]. So far, the WHO 1973 grading system has been the most widely accepted grading protocol among pathologists. The categorization of tumours relies on the complex interplay of various histopathological factors, doctors' observations and estimations on tissue structure and appearance. Accurate evaluation of histological material is mainly dependent on the pathologists' experience because different and complementary diagnostic variables and estimations are combined synergistically in assigning tumour grade. Taking into account the inherent subjectivity of any human decision, the final evaluation of tumour grading performed by histopathologists is questioned [3,4].

This work proposes an alternative methodology to develop a grade diagnostic tool, which is based on the formalization of specialized histopathological knowledge expressed in descriptive terms and concepts. The proposed method is based on fuzzy cognitive maps (FCMs) [5] and the utilization of active Hebbian learning (AHL) algorithm [6]. Our aim is to provide a new advanced diagnostic tool possessing three main characteristics: acceptable high diagnostic accuracy, transparency and interpretability.

This paper is structured in the following way: Section 2 reviews today's literature on computer-assisted methods for medical diagnosis, mainly on tumour classification and also outlines the main application areas of FCMs. Section 3 presents some basic aspects of FCM theory. Section 4 describes the methodology for developing the FCM model for tumour grading fuzzy cognitive map-grading tool (FCM-GT). Section 5 describes the FCM part of the grading process, how the decision regions for each grade category have been defined and how the proposed method is evaluated for 128 cases. Finally, the results of the tumour classification are compared with other techniques and discussed. Section 6 concludes the paper and gives future research directions.

2. Bibliography review

Many research efforts have been made to standardize the grading process of tumours which are mainly based either on machine learning techniques embedded in statistical classifiers or artificial neural networks (ANNs) [7–10]. Such computer-assisted methods have been used to increase the diagnostic and/or prognostic value of tumour grade classification.

Pattern recognition methods such as *k*-nearest neighbours, discriminant analysis, Bayesian classifiers (BNs), support vector machines (SVMs) and ANNs have been used for a number of disparate cancer diagnosis tasks: diagnosing breast cancer [11–13], prostate cancer [14], brain cancer [15,16], cervical cancer [17] and ovarian cancer [18].

Unlike previous methods, which are more data driven approaches, expert diagnosis systems incorporate prior knowledge and experience from experts' domain and they offer some insight as to how their diagnostic output derives. In recent years, fuzzy logic has been proved to be a powerful tool for diagnosis and decision-making systems. Fuzzy set theory and fuzzy logic are suitable tools for representing and handling imprecise and uncertain medical concepts and they have been successfully used to build medical expert systems [19].

As alternative methods to pattern classification, the fuzzy *k*-nearest algorithm and the neuro-fuzzy modelling (NFM) have been shown to be powerful soft pattern classifiers applied in medicine [20,21]. Soft classifiers such as the fuzzy *k*-nearest not only provide a diagnostic output but also indicate the degree to which the method is confident about its response. On the other hand, the NFM approach uses the modelling abilities of fuzzy logic to offer a degree of transparency in decision-making procedure.

Recent work has used image analysis methods for automatic grade classification employing quantitative tissue architectural and/or cytological features [7–10,22]. Specifically, researchers used linear discriminant analysis and tissue textural and/or struc-

tural features for the design of automatic grading systems [8,9]. However, these research groups relied on a subjective grading system according to which tumours are classified into four classes. Some other researchers made use of histological/cellular features estimated by histopathologists in conjunction with pattern recognition approaches [15,23]. Recently, Belacel and Boulassel introduced a fuzzy assigned method, the PROAFTN, for grading bladder tumours using features generated by image analysis [24].

The present study introduces a new diagnostic tool for assisting tumour grading. The proposed method lies in the cross-section of medical expert systems, soft computing and machine learning.

The proposed approach on assisting grade diagnosis takes advanced of fuzzy cognitive maps abilities. More specifically, FCM is a workable soft computing methodology that has been successfully applied in a number of discipline scientific areas. FCMs have been employed to model the causal inference [25], to make decision analysis in geographic information systems [26], to develop decision support systems [27], to perform failure modes and effects analysis in the process industry [28] and to model supervisory control systems [29,30].

In the medical application area, FCMs have been used to model and analyze the radiotherapy process and have been successfully used for decision-making in radiation therapy planning systems [31,32]. FCMs have also been proposed to analyze the problem of specific language impairment diagnosis in a critical way using several experts' opinions [33].

3. Fuzzy cognitive maps representation

The synergistic and complementary use of fuzzy logic and neuro-computing has initiated the development of soft computing methodologies. FCM is a soft computing technique that follows an approach similar to the human reasoning and human decision-making process. Soft computing methodologies have been investigated and proposed for the description and modelling of complex systems [34]. FCM consists of nodes (concepts) that illustrate the different aspects of the system's behaviour. These nodes (concepts) interact with each other showing the dynamics of the model. FCM is developed by human experts who operate/supervise/know the system and its behaviour under different circumstances in such a way that the accumulated experience and knowledge are integrated in a causal relationship between factors/characteristics/components of

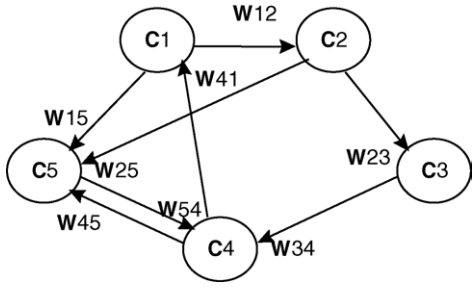

Figure 1 A simple fuzzy cognitive map representation.

the system [35]. Fig. 1 illustrates a graphical representation of a FCM.

Human knowledge and experience are reflected in the selection of concepts and weights for the interconnections between concepts of the FCM. Each node-concept represents one of the key-factors of the modelled system and it is characterized by a number A_i which represents its value. Concepts correspond to attributes, characteristics and qualities of the system. Interconnections among concepts of FCM signify the cause and effect relationship one concept has on the others. These weighted interconnections represent the direction and degree with which concepts influence the value of the interconnected concepts.

The cause and effect interconnection between two concepts C_j and C_i is described with the weight w_{ji}, taking value in the range -1 to 1.

There are three possible types of causal relationships between concepts:

- $w_{ji} > 0$: which indicates positive causality between concepts C_j and C_i. That is, an increase (decrease) in the value of C_j leads to an increase (decrease) in the value of C_i.
- $w_{ji} < 0$: which indicates negative causality between concepts C_j and C_i. That is, an increase (decrease) in the value of C_j leads to a decrease (increase) in the value of C_i.
- $w_{ji} = 0$: which indicates no relationship between C_j and C_i.

The value A_i of the concept C_i expresses the degree of its corresponding physical value. At each simulation step, the value A_i of a concept C_i is calculated by computing the influence of other concepts C_j's on the specific concept following the calculation rule:

$$A_i^{(k+1)} = f\left(A_i^{(k)} + \sum_{j\neq i, j=1}^{N} A_j^{(k)} \cdot w_{ji}\right) \quad (1)$$

where $A_i^{(k+1)}$ is the value of concept C_i at simulation step $k+1$, $A_j^{(k)}$ the value of concept C_j at simulation

step k, w_{ji} the weight of the interconnection from concept C_j to concept C_i and f is the sigmoid threshold function:

$$f = \frac{1}{1 + e^{-\lambda x}} \quad (2)$$

where $\lambda > 0$ is a parameter that determines its steepness. In this approach, the value $\lambda > 1$ has been used. This function is selected since the values A_i of the concepts lie within [0, 1].

The design of the FCM model is based on experts, using an interactive procedure of knowledge acquisition [29]. To increase objectivity in the development of FCMs, a group of experts is used. Experts are pooled together and determine the relevant factors that must be represented in the FCMs as concepts. They define the main concepts representing the model of the system on the basis of their knowledge and experience on the operation of the system. They know which factors are crucial and representative for the modelling of the medical system and they assign to each one a concept of FCM. Subsequently, they are separately asked to describe the relationship and the causality among concepts, using IF—THEN rules to justify the cause—effect relationship among concepts and to infer a linguistic weight for each interconnection. Every expert describes each one of the interconnections with a fuzzy rule; the inference of the rule is a linguistic variable which takes the values: "very very low" μ_{vvl}, "very low" μ_{vl}, "low" μ_{l}, "medium" μ_{m}, "high" μ_{h}, "very high" μ_{vh} and "very very high" μ_{vvh} (Fig. 2) [36]. Thus, there are assigned so many linguistic weights for one interconnection as the number of experts. Then, the inferred linguistic weights for each one interconnection are composed and an aggregated linguistic weight is produced using the fuzzy logic method SUM [37]. After that, the center of area (CoA) defuzzification method [37] is applied for the transformation of the overall linguistic weight to a numerical value for weight w_{ji}, belonging to the interval [−1, 1] and representing the overall suggestion of all the experts for this particular interconnection. Thus, an initial weight

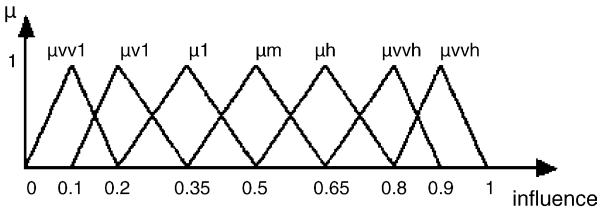

Figure 2 The seven membership functions corresponding to each one of the seven linguistic variables.

matrix, $\mathbf{W}^{initial} = [w_{ji}]$, i, $j = 1, \ldots, N$, with $w_{ji} = 0$, $i = 1, \ldots, N$, is obtained.

Experts use the fuzzy IF—THEN rule to describe the degree of influence from concept C_j to C_i, assumes the following form where **B**, **D** and **E** are fuzzy linguistic variables:

IF a change **B** occurs in the value of concept C_j, **THEN** a change **D** in the value of concept C_i is caused.

Infer: The influence from concept C_j to C_i is **E**.

The advantage of this methodology is that it is not required for experts to describe the causality relationships on numerical values, but rather to describe qualitatively the degree of causality among the concepts.

When the FCM model has been developed, the AHL algorithm is applied to adjusting the weights of the FCM interconnections and modifying them according to the specific problem characteristics. The AHL algorithm adapts all the weights of the FCM model using an acyclic fragment approach for concepts (asynchronous activation and interaction among concepts based on the initial experts' knowledge). The main advantage of the AHL algorithm is that it can determine new FCM causal links between all the concepts in order to increase classification capabilities of the FCM. In this way the AHL algorithm increases the FCMs' effectiveness, flexibility and robustness, and creates advanced FCMs with dynamic behaviour and great modelling abilities [6]. A detailed description and analysis of the AHL algorithm is provided in Appendix A.

4. Developing the FCM model for tumour grading

The FCM model for tumour grading was designed using the methodology described in Section 3. For this specific medical application, our experts were histopathologists with deep knowledge and great clinical experience, affiliated with the department of Pathology, University Hospital of Patras, Greece. Three experts defined the main histopathological features (concepts) and key characteristics, which encode the degree of tumour malignancy. These features are listed in Table 1 and are well documented in bibliography and represent the main variables that play an important role in the final grade diagnostic decision [38—41]. For this application, feature values take either two, three or four possible discrete or fuzzy values, as shown in Table 1.

Thus, experts designed a FCM model, which consists of nine concepts (Table 1). The eight concepts are the tumour features representing the main variables that histopathologists usually take into

Table 1 Histological features for coding tumours' malignancy

Histological features	Assessment	Type of values scaled
C_1: cell distribution	Even, clustered	Two discrete (0 or 1)
C_2: cell size	Uniform, pleomorphic	Two discrete (0 or 1)
C_3: cell number	Numerous, variable	Two discrete (0 or 1)
C_4: cytoplasm	Homogeneous, variable	Two discrete (0 or 1)
C_5: nuclei	Uniform, irregular, very irregular, bizarre	Four fuzzy values (zero, low, medium, high)
C_6: nucleoli	Inconspicuous, evident, prominent	Three discrete (0, 0.5 or 1)
C_7: necrosis	Inconspicuous, frequent	Two discrete (0 or 1)
C_8: mitosis	Absent rate, occasional, numerous	Three discrete (0, 0.5 or 1)

consideration in assigning grade to the bladder tumours. The ninth concept represents the tumour grade. More specifically, concept C_1 represents the cell distribution, C_2 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 represents the tumour grade.

Three histopathologists experts determined the concepts of FCM. Each expert was asked to define the degree of influence among the concepts and describe their interrelationship using an IF—THEN rule. Then, experts inferred a linguistic weight to describe the cause and effect relationship between every pair of concepts.

The three histopathologists experts suggested that the degree of influence between concepts was described by a linguistic variable taking value in [0, 1] and its fuzzy set was defined as: T(influence) = {very very low, very low, low, medium, high, very high, very very high} [34]. Each element of the fuzzy set corresponds to a membership function shown in Fig. 2. It is noticeable that these membership functions have a finer distinction between grades in the lowest and highest end of the influence scale.

Some examples of the fuzzy rules that experts proposed are given:

The following rule describes the influence of concept C_1 (cell distribution) on concept C_9 (tumour grade):

IF a *small* change occurs in the value of concept C_1 (cell distribution), **THEN** a *small* change is caused in the value of concept C_9 (tumour grade).

Infer: The influence from C_1 to C_9 is *positive very low*.

Another rule presents the influence from concept C_5 (nuclei) towards concept C_8 (mitosis):

IF a *small* change occurs in the value of concept C_5, **THEN** a *large* change is caused in the value of C_8.

Infer: The influence from C_5 to C_8 is *positive high*.

This means that if the type of nuclei changes, the value of mitosis increases very much.

The inferred fuzzy linguistic variables for each interconnection — determined by the group of three experts — are combined and thus an aggregated linguistic weight is produced using the SUM method, which is then deffuzified with the CoA method [27]. The result is a crisp value representing the weight for each interconnection.

To illustrate how numerical values of weights are produced, the experts' suggestions on how to indicate the interconnection between concept C_5 (nuclei) and concept C_9 (tumour grade) are shown below:

1st expert:

IF a *small* change occurs in the value of concept C_5, **THEN** a *medium* change in value of concept C_9 is caused.

Infer: The influence from C_5 to C_9 is *positive medium*.

2nd expert:

IF a *small* change occurs in the value of concept C_5, **THEN** a *large* change in value of concept C_9 is caused.

Infer: The influence from C_5 to C_9 is *positive high*.

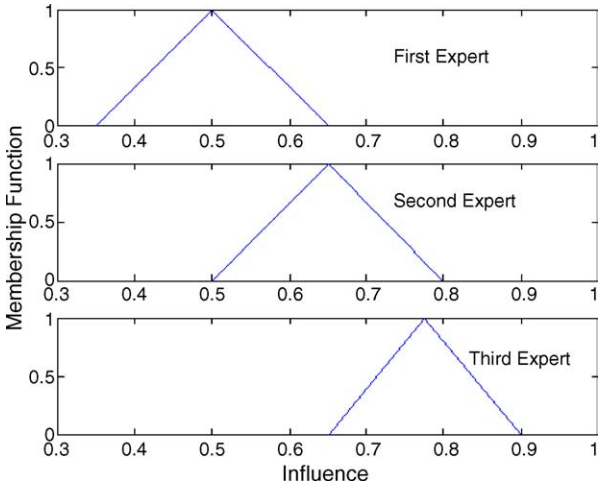

Figure 3 Example of three linguistic variables suggested by the three experts to describe the relationship between two concepts.

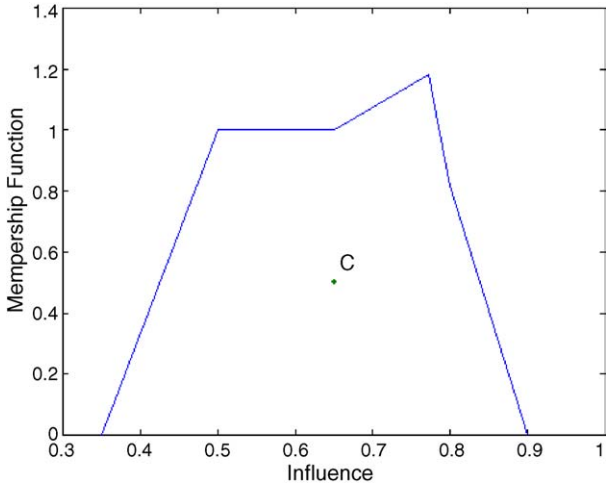

Figure 4 Aggregation of three linguistic variables using the SUM technique. Point C is the numerical weight after defuzzification using the CoA method.

3rd expert:

IF a *very small* change occurs in the value of concept C_5, **THEN** a *large* change in value of concept C_9 is caused.

Infer: The influence from C_5 to C_9 is *positive very high*.

Fig. 3 illustrates the three suggested linguistic variables, for this particular example.

These linguistic variables (medium, positive high and positive very high) are summed and an overall linguistic weight is produced (Fig. 4), which with the defuzzification method of CoA is transformed into the numerical value of $w_{59} = 0.65$.

The same approach was used to determine all the weights of the FCM. A weight matrix $\mathbf{W}^{\text{initial}}$ gathering the initially suggested weights of all the interconnections among the concepts of the FCM model was produced. Fig. 5 illustrates the FCM tumour grading model (FCM-GT), consisting of 9 concepts and 21 weighted interconnections.

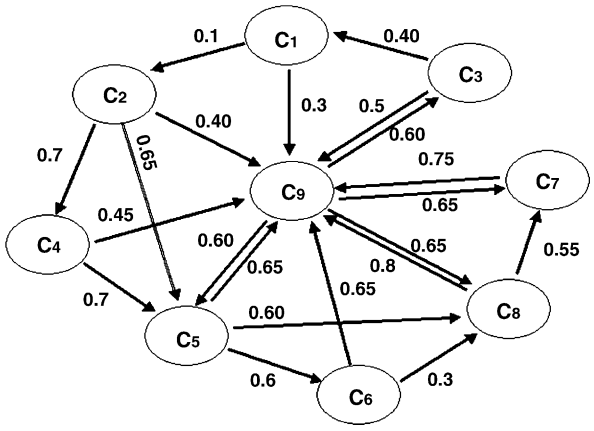

Figure 5 The FCM tumour grading model.

When the FCM-GT has been developed it could be used to assign grade of every tumour. But for better classification results, the unsupervised learning algorithm of AHL is applied. The AHL algorithm modifies the weights of the FCM model, according to each case, so as to ensure that the FCM-GT determines successfully the value of grade concept C_9, which indicates the category of tumour. The implementation of the AHL algorithm requires that the three experts had selected the Activation and Activated concepts and the sequence of activation [6]. The ninth concept, which determines the tumour grade, was defined as the activation decision concept (ADC). Experts described the sequence of activation and thus the concepts "mitosis" (C_8) and "necrosis" (C_7) were defined as the first activation concepts. Concepts "cell distribution" (C_1), "cell size" (C_2), "cell number" (C_3) and "cytoplasm" (C_4) behave as a second set of activation concepts at the next step. The concepts "nuclei" (C_5) and "nucleoli" (C_8) are the third set of activation concepts. All concepts together influence concept C_9, whose value is calculated using Eq. (A.4). The suggested sequence of activated and activation concepts (Fig. 6) are in accordance with the way in which the three histopathologists experts examine the histological material microscopically in order to

$$\mathbf{W}^{\text{initial}} = \begin{bmatrix} 0 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 \\ 0 & 0 & 0 & 0.7 & 0.65 & 0 & 0 & 0 & 0.40 \\ 0.4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.50 \\ 0 & 0 & 0 & 0 & 0.7 & 0 & 0 & 0 & 0.45 \\ 0 & 0 & 0 & 0 & 0 & 0.6 & 0 & 0.60 & 0.65 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3 & 0.65 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.75 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.55 & 0 & 0.80 \\ 0 & 0 & 0.60 & 0 & 0.60 & 0 & 0.65 & 0.65 & 0 \end{bmatrix}$$

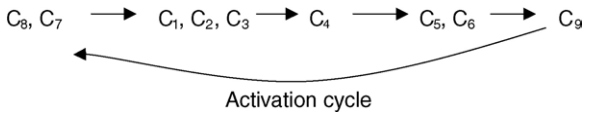

Figure 6 The sequence of activation.

assign the grade of tumour. Histopathologists start by 'scanning' the tissue sample under the microscope in order to assess the tissue appearance as a whole (examining "mitosis" and "necrosis"). Then, they assess morphological nuclear features (examining cell "distribution", "size" and "number"), they focus on regions with marked nuclear atypia (examining "nuclei" and "nucleoli") and finally they conclude about tumour grade.

5. Novel method for tumour grading

5.1. The fuzzy cognitive map part of the grading process

When the FCM-GT has been designed and the necessary specifications for the implementation of the AHL algorithm have determined, the FCM-GT was tested to evaluate 128 cases (tissue biopsies) of urinary bladder cancer.

The data-cases (specimens) of urinary bladder cancer were retrieved from the archives of the Department of Histopathology of the University Hospital of Patras, Greece. Tissue sections of tumours were fixed and stained according to the standard Hematoxylin—Eosin method [2]. Three histopathologists experts reviewed a great number of specimens but only those cases were used for which the three histopathologists were in full agreement regarding tumour characterization and thus only 128 cases were selected. Following the conventional WHO grading system [42], experts classified unanimously the cases as follows: 40 cases as grade I, 43 as grade II and 45 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). These values were transformed in the range [0, 1], and were assigned to the corresponding concepts of the FCM grading model.

The following mathematical form was used to transform the quantitative values of each one of the eight characteristics to a numerical value [15]:

$$\text{Submitted value} = \frac{[\text{option selected}] - 1}{[\text{total number of options}] - 1} \tag{3}$$

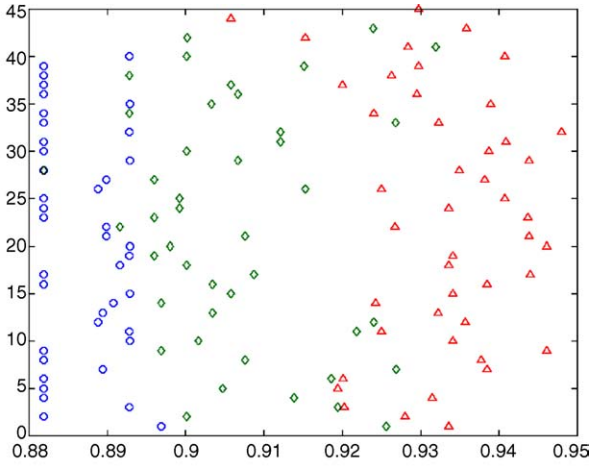

Figure 7 The estimated 'Grade' values for the 128 cases.

For example, let us say that one histopathologist describes feature "nuclei" (concept C_5) as "irregular", which is the second of four possible options (e.g. "uniform" "irregular" "very irregular" and "bizarre"), so the assigned value would be:

$$\frac{[\text{option 2 selected}] - 1}{[\text{total of 4 options}] - 1} = \frac{2-1}{4-1} = \frac{1}{3} \tag{4}$$

The initial value of concept C_9 was set a random value in [0, 1], which was the same for all cases. Then, for each case, the FCM-GT with the initial values of concepts and with the initial weight matrix $\mathbf{W}^{\text{initial}}$ starts to examine the grade of tumour. The AHL algorithm is employed for every case to modify the FCM weights and finally the value of concept C_9 is calculated, which is the grade for this case.

Fig. 7 illustrates the values of concept C_9 (we refer to this value as "Grade") for each of the 128 cases. For grade I cases, the estimated "Grade" values are represented by '○'; for grade II cases, the estimated "Grade" values are represented by '□'; for grade III cases, the estimated "Grade" values are represented by '▽'.

5.2. Definition of decision boundaries and final classification stage

Then the decision regions associated with each tumour grade category are determined (grade I, II or III). For this purpose, one-dimensional decision boundaries were determined employing the Bayesian statistical decision method [43].

Let G_1, G_2 and G_3 be the classes, which contain the estimated "Grade" values for grades I, II and III cases, respectively.

According to Bayesian method an estimated value G — for an unknown case — is assigned one of the

most probable classes (G_1, G_2 or G_3) or to the class with the greater conditional probability $P(G_i|G)$.

The probability $P(G_i|G)$ (Bayes rule) is defined as:

$$P(G_i|G) = \frac{P(G|G_i)P(G_i)}{P(G)} \quad (5)$$

where

- $P(G)$ is the probability density function (pdf) of values G:

$$P(G) = \sum_{i=1}^{3} P(G|G_i) \times P(G_i) \quad (6)$$

- $P(G|G_i)$ is the grade class-conditional probability density function, describing the distribution of the grade values G in each one of the grade classes G_i, $i = 1-3$. Assuming that $P(G|G_i)$ follows a Gaussian distribution, the value of:

$$P(G|G_i) = \frac{1}{\sigma_i \,\mathrm{sqrt}(2\pi)} \exp\left(-\frac{1}{2}\left(\frac{G - m_i}{\sigma_i}\right)^2\right) \quad (7)$$

  can be easily estimated using the available data, σ_i is the standard deviation and m_i is the mean value of grade values G for the cases belonging to grade class G_i.
- $P(G_i)$ is the a priori probability estimated from the available data: assuming N is the total number of cases and N_i of them belonging to G_i then the a priori probability for each class G_i, is: $P(G_i) \approx N_i/N$.

The Bayesian classifier is designed to classify an unknown case as follows:

- If $\max\{P(G_1|G), P(G_2|G), P(G_3|G)\}$ is $P(G_1|G)$, then the case with grade value G is classified as tumour with grade I (class G_1).
- If $\max\{P(G_1|G), P(G_2|G), P(G_3|G)\}$ is $P(G_2|G)$, then the case with grade value G is classified as tumour with grade II (class G_2).
- If $\max\{P(G_1|G), P(G_2|G), P(G_3|G)\}$ is $P(G_3|G)$, then the case with grade value G is classified as tumour with grade III (class G_3).

To compute optimal decision boundary between G_1 and G_2 classes, we equate their posterior probabilities:

$$P(G_1|G) = P(G_2|G) \quad (8)$$

Similarly to compute decision boundary between G_2 and G_3 classes, we equate:

$$P(G_2|G) = P(G_3|G) \quad (9)$$

Choosing the most likely class for each case we minimize the overall error rate. However, in this particular medical application the eventual risk of misclassifying a high-grade case, as grade II (or I), is much higher than the reciprocal error. Also the risk for misclassifying a grade II case as grade I is higher than the opposite. To satisfy these requirements in estimating the decision boundaries, the eventual risk of misclassifying a high grade case and the risk of misclassifying a grade II case as grade I were taken into consideration; the posterior probabilities were multiplied by penalty terms (see for a more formal description [43]), and new decision boundaries that minimize the overall risk were estimated as follows:

$$\lambda_{12} \times P(G_1|G) = \lambda_{21} \times P(G_2|G) \quad (10)$$

where λ_{12} is the penalty term for misclassifying grade I case as grade II and λ_{21} is the penalty term for misclassifying grade II as grade I, and

$$\lambda_{23} \times P(G_2|G) = \lambda_{32} \times P(G_3|G) \quad (11)$$

where λ_{23} is the penalty term for misclassifying grade II case as grade III and λ_{32} is the penalty term for misclassifying grade III case as grade II.

Penalty terms were adjusted so as to achieve the desired trade off between correct diagnosed grades III and II cases; correct diagnosed grades II and I cases. More specifically, $\lambda_{12} = 0.1$, $\lambda_{21} = 0.3$, $\lambda_{23} = 0.3$ and $\lambda_{32} = 1$.

Fig. 8 shows the estimated one-dimensional decision boundaries following the pre-described Bayesian approach. The decision boundaries are equal to 0.891 and 0.918, respectively.

To evaluate the accuracy with which cases were classified as grade I, II or III, the leave one out method was employed. According to this method, each time the decision boundaries were constructed using all

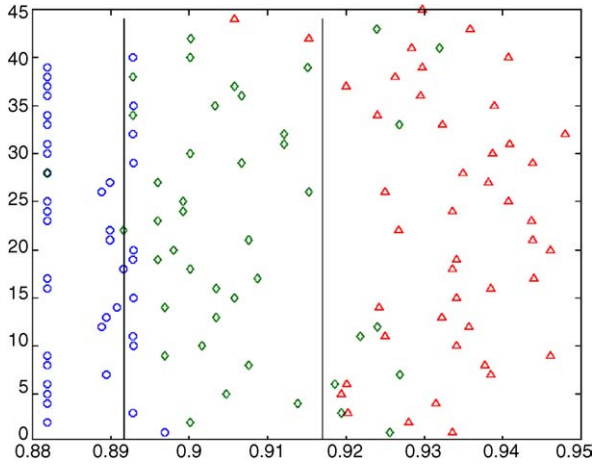

Figure 8 The classification of the 128 cases along with the decision boundaries.

Table 2 Truth table of the FCM-based tumour grading process

Histological finding	FCM-based tumour grading process (FCM-GT)			
	Grade I	Grade II	Grade III	Accuracy (%)
Grade I	29	11	0	72.55
Grade II	2	32	9	74.42
Grade III	0	2	43	95.55

cases but one, which was then used as a test case. The leave one out method was repeated for all cases of each grade category and the results were presented in a truth table (Table 2) revealing the satisfying accuracy of the tool; the accuracy for grades I, II and III cases was: 72.5% (29/40), 74.42% (32/43) and 95.55% (43/45), respectively. More specifically, these accuracies percentages provide the degree of confidence with which we can rely on the FCM grading outcome for a given new case.

5.3. Discussion of results

The proposed method based on FCMs for tumour grading provides a framework within which histopathologists evaluate a series of traditional diagnostic concepts (features). The way the FCM grading model is designed increases the objectivity of the diagnostic process by taking into account the different experts' opinions regarding the interplay of histopathological variables in the final grade diagnostic output. Using these variables, the FCM model estimates a grade value according to which a particular case is classified as grade I, II or III.

Concerning the grade diagnostic accuracy, the methodology exhibited a significant high sensitivity of 95.55% for the high-risk (grade III) tumours. A lower accuracy of 72.55 and 74.42% was obtained for tumours: grades I and II, respectively. However, in the existing literature, similar or lower levels of confidence have been reported with which histopathologists diagnose urinary bladder cancers of: grade I or II [3]. It is also worth mentioning that no case of grade III was misclassified as grade I and vice versa. These results suggest an accepted agreement between the proposed method and grade assigned by the histopathologists.

In recent work, exactly the same histopathological set of data (Table 1) was fed directly to a multilayer neural network — employing the back-propagation learning algorithm — to diagnose tumours' grade and the following classification results were obtained: 64.9% for grade I, 69.4% for grade II and 82.7% for grade III [22,44].

Other studies on computer-aided grade classification have relied on pattern classification approaches. Choi et al. [8] and Jarkans et al. [9] have used linear discriminant analysis and tissue textural and/or structural features for the design of automatic grading systems. But, both researchers have relied on a different subjective grading system (a modification of the 1973 grading protocol) according to which tumours are classified into four classes (grades I, IIa, IIb and III).

In another work, researchers have proposed a computer-based grading system using ANNs in conjunction with quantitative cell nuclei features and subjectively evaluated histopathological variables. Their system classified tumours of grades I, II and III with an accuracy of 82, 80.5 and 93.1%, respectively [23].

In a more recent work, Belacel et al. have applied a multicriteria fuzzy assignment method PROAFTN, and a panel of 24 quantitative parameters derived from computer-assisted microscopy, to perform tumour grading. The method exhibited the following accuracies: 71.3% for grade I, 52% for grade II and 57.7% for grade III tumours [24].

Table 3 gathers all the results from the three different techniques.

Whereas the accuracy of the FCM model is very important it possesses several other benefits in comparison to other pattern classification approaches, which make FCM-GT more acceptable to the clinical practice. It offers a degree of transparency on diagnostic knowledge and FCM-GT has the ability to explain decisions accurately when diagnosing new cases. Because it is very important for a physician to be able to analyze and understand how the method derives the diagnostic output. In this way, the proposed model would be beneficial to both experts and practising histopathologists. For the former, the model is capable of revealing new interrelations and regularities that might not see before in an explicit way providing thus the experts with a novel point of view on the specific problem of

Table 3 Comparison of three different techniques applying to same data

Histopathological finding	FCM-GT (%)	NNs (back-propagation) (%)	PROAFTN (%)
Grade I	72.55	64.9	71.13
Grade II	74.42	69.4	52
Grade III	95.55	82.7	57.73

grading tumours. For the latter, the FCM model could be used for training purposes since not only it highlights the areas where diagnostic criteria seem to be vague but also it offers an explanation of the various variables interplay in the final outcome.

Finally, the proposed methodology offers a flexible modelling method where new features can easily be introduced, added or deleted in the grade model following histopathologists grading criteria that continuously evolve.

6. Summary and closing remarks

In this paper, a novel advanced method to support grade diagnostic decision is proposed and analyzed. The novelty of the method is based on employing the FCMs to represent specialized knowledge (experience, expertise and heuristic) on histopathology and on using the AHL algorithm. The proposed FCM-GT proves to be an efficient and dynamic model for automatic grade characterization. The proposed method works with reasonably high accuracy compared to other existing methods and at the same time fulfils the physicians' requirements for transparency and explicability.

Proposed future research work could be directed towards the development of an integrated two-level hierarchical structure. In the lower level, values of some histopathological variables could be automatically extracted from image analysis methods. These values will be fed to the upper-level where an advanced FCM-GT tool accomplished with new concepts will perform automatic grade classification.

Acknowledgements

This work was supported by the ''Pythagoras II'' research grant co-funded by the European Social Fund and National Resources.

Authors would like to thank the histopathologists: Ravazoula Panagiota, Batistatou Anna and Aletra Christina from the Department of Pathology, University Hospital of Patras, for their suggestions and help in developing the FCM model and for providing us with the clinical data.

Appendix A

The AHL algorithm introduces the asynchronous updating for the concepts and weights of FCMs [6]. It requires the definition of the activation and activated concepts, the sequence of activation as well as the activation decision concepts (ADCs), which are the observable outputs of the system. The ADCs concepts are our main concern and we want to estimate their values. The AHL algorithm is based on the premise that there are some steps of activations (sequence of activation). A simulation cycle consists of steps, at each activation step one or more concepts are the activation concepts that influence the interconnected ones and so on until the termination of the sequence of activation that close the cycle. This concept, at the next iteration step, becomes activated concept. For example, let us say the jth concept C_j, is the triggering concept that influences concept C_j. The concept C_j is declared the activation concept, with the value A_j^{act} and it triggers the interconnected corresponding concept C_i, which is the activated concept. At the next iteration step, the concept C_i influences the other interconnected concepts C_i and so on. It is assumed that there is asynchronous stimulation mode due to which the concept C_l becomes the activation concept that triggers C_l and the other interconnected concepts and there is a sequence of activation steps. During this activation process, the weight w_{ji} of the causal interconnection of the related concepts is updated and the modified weight $w_{ji}^{(k)}$ is derived for each iteration step k.

It should be noticed here that experts initially determined the activated and activation concepts for every activation step, according to the infrastructure of the FCM and the system itself. So, the sequence of activation and activated concepts determine the way with which factors-concepts affect the ADCs [6].

The AHL adjusts the weights between concepts at each activation step using the following discrete type of asynchronous mode:

$$w_{ji}(k) = \gamma \cdot w_{ji}(k-1) + \eta \cdot A_j^{\text{act}}(k-1) \cdot A_i(k-1) \tag{A.1}$$

where the η is the learning rate parameter and γ is the weight decay parameter. Here, it is supposed that for the activation step k, concept C_j with value A_j^{act} is the activated concept and concept C_i with value A_i is the interconnected activation concept at the same simulation step. The parameters η and γ take positive values and it is also supposed that $\gamma > \eta$.

The learning rate parameter η is a small positive scalar parameter that is defined to decrease exponentially with activation cycle c, following the equation:

$$\eta^{(c)} = b_1 \cdot \exp(-\lambda_1 \cdot c) \tag{A.2}$$

Behaviour of FCMs depends on the step size $\eta^{(c)}$ decay with time, thus $\eta^{(c)}$ is selected to decrease and the rate of decrease depends on the speed of convergence to the optimum solution and on the activation mode. The parameters b_1 and λ_1 are positive learning factors, which are determined using the trial and error method procedure so that to optimize the final solution.

The learning factor $\eta^{(c)}$ takes the following values that ensure fast convergence of concepts values:

$$\eta^{(c)} = 0.02 \cdot \exp(-0.2 \cdot c) \tag{A.3}$$

The weight decay coefficient γ may be zero, constant or may decrease by the number of iteration steps c, this depends on the problem's constraints. The parameter γ takes the constant value of 0.98 for this specific problem to ensure that the learning process converges in a desired steady state.

Eq. (1) that calculates the value of each concept of FCM is updating, taking the form of Eq. (A.4) where the value of weight $w_{ji}^{(k)}$ is calculated using Eq. (A.1):

$$A_i(k+1) = f\left(A_i(k) + \sum_{j \neq i, j=1}^{N} A_j^{\text{act}}(k) \cdot w_{ji}(k)\right) \tag{A.4}$$

The FCM simulation model is now based on the repetitive multiplication of the activated concept with value A_j^{act} at iteration number k, with the updated connection weight $w_{ji}(k)$.

References

[1] Bostwick D, Ramnani D, Cheng L. Diagnosis and grading of bladder cancer and associated lesions. Urol Clin N Am 1999;26:493—507.

[2] Mostofi F, Sobin L, Torloni H. Histological typing of urinary bladder tumours, vol. 10. Geneva: World Health Organization, 1973.

[3] Ooms E, Anderson W, Alons C, Boon M, Veldhuizen R. Analysis of the performance of pathologists in grading of bladder tumours. Hum Pathol 1983;14:140—3.

[4] Witjes J, Kiemene L, Wheeless L, Grossman H. The value of histopathological prognostic factors in superficial bladder cancer: do we need more? Urol Oncol 2000;5:185—90.

[5] Kosko B. Fuzzy cognitive maps. Int J Man Mach Stud 1986;24:65—75.

[6] Papageorgiou E, Stylios C, Groumpos P. Active Hebbian learning algorithm to train fuzzy cognitive maps. Int J Approx Reason 2004;37(3):145—253.

[7] Van der Poel H, Schaafsma H, Vooijs P, Debruyne F, Schalken J. Quantitative light microscopy in urological oncology. Review article. J Urol 1992;148:1—13.

[8] Choi H, Vasko J, Bengtsson E, Jarkrans T, Malmstrom U, Wester K, et al. Grading of transitional cell bladder carcinoma by texture analysis of histological sections. Anal Cell Pathol 1994;6:327—43.

[9] Jarkrans T, Vasko J, Bengtsson E, Choi H, Malmstrom U, Wester K, et al. Grading of transitional cell bladder carcinoma by image analysis of histological sections. Anal Cell Pathol 1995;18:135—58.

[10] Spyridonos P, Ravazoula P, Cavouras D, Berberidis K, Nikiforidis G. Computer-based grading of haematoxylin—eosin stained tissue sections of urinary bladder carcinomas. Med Inform Internet Med 2001;26(3):179—90.

[11] Finette S, Bleier A, Swindell W. Breast tissue classification using diagnostic ultrasound and pattern recognition techniques: I. Methods of pattern recognition. Ultrason Imag 1983;5:55—70.

[12] Wolberg W, Street W, Mangasarian O. Image analysis and machine learning applied to breast cancer diagnosis and prognosis. Anal Quant Cytol Histol 1995;17(2):77—87.

[13] Cheng H, Cai X, Chen X, Hu L, Lou X. Computer-aided detection and classification of microcalsifications in mammograms: a survey. Pattern Recognit 2003;16(12):2967—91.

[14] Anagnostou T, Mesut R, Lykourinas M, Djavan B. Artificial neural networks for decision making in urologic oncology. Eur Urol 2003;43(6):596—603.

[15] Mckeon J, Ramsay D. Classification of astrocytomas and malignant astrocytomas by principal components analysis and a neural net. J Neuropath Exp Neurol 1996;55:1238—45.

[16] Glotsos D, Spyridonos P, Petalas P, Cavouras D, Ravazoula P, Dadioti P, et al. 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 Histol Cytol 2004;26(2):77—83.

[17] Timmerman D. The use of mathematical models to evaluate pelvic masses; can they beat an expert operator? Best practice & research. Clin Obstet Gynecol 2004;18(1):91—104.

[18] Lu C, Van Gestel T, Suykens J, Van Huffel S, Vergote I, Timmerman D. Preoperative prediction of malignancy of ovarian tumours using least squares support vector machines. Artif Intell Med 2003;28(3):281—306.

[19] Nguyen H, Vladik K. Fuzzy logic and its application in medicine. Int J Med Inf 2001;62:165—73.

[20] Seker H, Odetayo M, Petrovic D, Naguib R. A fuzzy logic based-method for prognostic decision making in breast and prostate cancers. IEEE Trans Inform Technol Biomed 2003;7(2):114—22.

[21] Catto J, Linkens D, Abbod M, Chen M, Burton J, Feeley K, et al. Artificial intelligence in predicting bladder cancer outcome: a comparison of neuro-fuzzy modeling and artificial neural networks. Artif Intell Med 2003;9:4172—7.

[22] Spyridonos P, Cavouras D, Ravazoula P, Nikiforidis G. Neural network based segmentation and classification system for the automatic grading of histological sections of urinary bladder carcinoma. Anal Quant Cytol Histol 2002; 24(6):317—24.

[23] Spyridonos P, Cavouras D, Ravazoula P, Nikiforidis G. A computer-based diagnostic and prognostic system for assessing urinary bladder tumour grade and predicting cancer recurrence. Med Inform Internet Med 2002;27(2):111—22.

[24] Belacel N, Boulassel M. Multicriteria fuzzy assignment method: a useful tool to assist medical diagnosis. Artif Intell Med 2001;21:201—7.

[25] Miao Y, Liu Z. On causal inference in fuzzy cognitive maps. IEEE Trans Fuzz Syst 2000;8:107—19.

[26] Liu Z, Satur R. Contextual fuzzy cognitive map for decision support in geographic information systems. IEEE Trans Fuzz Syst 1999;5:495—507.

[27] Khan S, Chong A, Gedeon TA. A methodology for developing adaptive fuzzy cognitive maps for decision support. J Adv Comput Intell 2000;4(6):403—7.
[28] Pelaez C, Bowles B. Using fuzzy cognitive maps as a system model for failure modes and effects analysis. Inform Sci 1996;88:177—99.
[29] Stylios C, Groumpos P, Georgopoulos V. A fuzzy cognitive maps approach to process control systems. J Adv Comput Intell 1999;3(5):409—17.
[30] Groumpos P, Stylios C. Modelling supervisory control systems using fuzzy cognitive maps. Chaos Solit Fract 2000;11(1—3):329—36.
[31] Papageorgiou E, Stylios C, Groumpos P. Decision Making in external beam radiation therapy based on FCMs. In: Proceedings of the 1st international IEEE symposium on intelligent systems; 2002 [CD-ROM].
[32] Papageorgiou E, Stylios C, Groumpos P. An integrated two-level hierarchical decision making system based on fuzzy cognitive maps (FCMs). IEEE Trans Biomed Eng 2003; 50(12):1326—39.
[33] Georgopoulos V, Malandraki G, Stylios C. A fuzzy cognitive map approach to differential diagnosis of specific language impairment. Artif Intell Med 2002;29(3):261—78.
[34] Jang J, Sun C, Mizutani E. Neuro-fuzzy and soft computing. NJ: Prentice-Hall, 1997.
[35] Kosko B. Neural networks and fuzzy systems. Englewood Cliffs, NJ: Prentice-Hall, 1992.
[36] Stylios C, Groumpos P. Fuzzy cognitive maps in modelling supervisory control systems. J Intell Fuzz Syst 2000;8:83—98.
[37] Lin C, Lee C. Neural fuzzy systems: a neuro-fuzzy synergism to intelligent systems. NJ: Prentice Hall, 1996.
[38] Murphy W, Soloway S, Jukkola A, Crabtree W, Ford K. Urinary cytology and bladder cancer. The cellular features of transitional cell neoplasms. Cancer 1984;53:1555—65.
[39] Murphy W. Urothelial neoplasia. Pathology and pathobiology of the urinary bladder and prostate. Baltimore: Williams & Wilkins, 1992.
[40] Carbin E, Ekman P, Gustafson H, Christensen N, Sandstedt B, Silfversward C. Grading of human urothelial carcinoma based on nuclear atypia and mitotic frequency. Part I. Histological description. J Urol 1998;6:968—71.
[41] Bostwick D, Mikuz G. Urothelial papillary neoplasms. Virchows Arch 2002;441:109—16.
[42] Mostofi K, Davis CJ, Sesterhenn IA. WHO histologic typing of urinary bladder tumours. Berlin: Springer, 1999.
[43] Theodoridis S, Koutroubas K. Classifiers based on Bayes decision theory. In: Pattern recognition2nd ed., USA: Elsevier Science/Academic Press, 2003. p. 13—44.
[44] Spyridonos P. Computer-aided grading of tumours using image analysis and supervised learning algorithms. Ph.D. thesis. Greece: Computer Laboratory, School of Medicine, University of Patras; 2002.