

An image-analysis system based on support vector machines for automatic grade diagnosis of brain-tumour astrocytomas in clinical routine

D. GLOTSOS¹, P. SPYRIDONOS¹, D. CAVOURAS², P. RAVAZOULA³,
P. ARAPANTONI DADIOTI⁴ & G. NIKIFORIDIS¹

¹Medical Image Processing and Analysis Laboratory, School of Medicine, University of Patras, Plo, Patras, 265 00, Greece ²Department of Medical Instrumentation Technology, Technological Educational Institution of Athens, Ag. Spyridonos Street, Aigaleo, 122 10 Athens, Greece ³Department of Pathology, University Hospital, Rio, Patras, 265 00, Greece ⁴Department of Pathology, General Anticancer METAXA Hospital, Pireas, 185 37, Greece

(Received February 2004; accepted February 2005)

Abstract

An image-analysis system based on the concept of Support Vector Machines (SVM) was developed to assist in grade diagnosis of brain tumour astrocytomas in clinical routine. One hundred and forty biopsies of astrocytomas were characterized according to the WHO system as grade II, III and IV. Images from biopsies were digitized, and cell nuclei regions were automatically detected by encoding texture variations in a set of wavelet, autocorrelation and parzen estimated descriptors and using an unsupervised SVM clustering methodology. Based on morphological and textural nuclear features, a decision-tree classification scheme distinguished between different grades of tumours employing an SVM classifier. The system was validated for clinical material collected from two different hospitals. On average, the SVM clustering algorithm correctly identified and accurately delineated 95% of all nuclei. Low-grade tumours were distinguished from high-grade tumours with an accuracy of 90.2% and grade III from grade IV with an accuracy of 88.3%. The system was tested in a new clinical data set, and the classification rates were 87.5 and 83.8%, respectively. Segmentation and classification results are very encouraging, considering that the method was developed based on every-day clinical standards. The proposed methodology might be used in parallel with conventional grading to support the regular diagnostic procedure and reduce subjectivity in astrocytomas grading.

Keywords: *Support vector machines; astrocytomas; grading; wavelets*

1. Introduction

In diagnosing brain astrocytomas, the determination of the degree of tumour abnormality (grade) is of major clinical importance because it specifies treatment planning and patient management [1]. To determine the grade of a tumour, the pathologist examines biopsy material under a microscope. According to the WHO grading system, astrocytomas are

classified into prognostic categories: low grade (grade II), and high grade (grade III and grade IV), with grade IV tumours being the most aggressive type [2]. However, the subjectivity in interpretation and physicians' experience have been shown to significantly influence the quality of diagnosis. Inter- and intra-observer accuracy can be as low as 40% [3]. These variations promote sources of diagnostic errors, which may lead to false treatment planning with adverse effects in patient survival [4].

To account for physician's subjectivity, many studies [5–11] proposed computer-aided systems for objectifying the diagnostic process. However, most of these approaches have suggested automatic diagnostic methods, based on specialized clinical protocols, which cannot be easily applied in daily clinical practice [5–11]. Furthermore, these studies have evaluated mostly the separation of low- from high-grade tumours [5–8], since until recently, the belief had been that grade III and grade IV tumours have to be treated identically [1]. New published data, though, have indicated that grade III tumours are more chemosensitive than grade IV [12–13]. It would be thus very essential to construct a grading system that will optimally discriminate not only low from high but also grade III from grade IV tumours, based on every-day clinical standards; such standards are considered to be the WHO grading system and the Hematoxylin-Eosin (H&E) staining protocol [3,7].

However, in order to construct an automatic grading approach compatible with clinical routine, there must be a reliable and reproducible way to process and segment H&E-stained images. Literature has highlighted the difficulty in processing this kind of histological images: the segmentation task is not straightforward due to the diversity in texture among nuclei, surrounding tissue and other existing structures [14,15]. Consequently, mostly specialized staining procedures have been examined such as the Feulgen and ki-67 staining [5–11]. In Feulgen and ki-67-stained images, nuclei are well distinguishable from the background, but the preparation processes are much more complicated and time-consuming, and cannot be easily applied in daily clinical practice.

In this study, we propose a computer-based image analysis system based on the concept of Support Vector Machines for grading brain astrocytomas in clinical routine. H&E images were automatically segmented for the separation of nuclei from surrounding tissue by encoding texture variations in a set of wavelet, autocorrelation and parzen estimated descriptors and using an unsupervised SVM clustering methodology. Based on morphological and textural nuclear features, a decision-tree classification scheme discriminated between different grades of tumours employing an SVM classifier. To the best of our knowledge, there is no approach for texture-based unsupervised segmentation of any type of histological images; furthermore, SVM-clustering is a state-of-the-art method, which until now had not been examined beyond its theoretical foundation [16]. Finally, in contrast to previous studies [5–11] that have constructed their automatic grading systems with data collected from the same hospital [5–11], we evaluated the generalization of our method validating clinical material collected from two different hospitals.

2. Material and methods

2.1. Material

The clinical material comprised 140 biopsies of astrocytomas collected from the Departments of Pathology of the University Hospital of Patras, Greece (92 cases) and the General Distinct Anticancer Hospital of Piraeus METAXA, Greece (48 cases). Five H&E-stained sections were generated from the same block (patient). Two independent pathologists, one from each hospital, reviewed each group of slides separately under the microscope. In case of discrepant

diagnosis, a consensus meeting was held. Of the 140 astrocytomas, 61 were classified as low grade (GII) and 79 as high grade (30 GIII and 49 GIV) according to the WHO grading system. For each slide, a pathologist specified the most representative region. From this region, images have been acquired ($768 \times 576 \times 8$ bit) using a light Zeiss Axiostar plus microscope connected to an Ikegami colour video camera. Figure 1 depicts a typical H&E-stained image of astrocytomas.

2.2. Segmentation of H&E-stained images of astrocytomas

2.2.1. Design of the segmentation algorithm. Each image pixel was described with five textural features extracted from the autocorrelation function, wavelets and a density estimation parzen motivated method. Based on this feature vector, an SVM-clustering algorithm classified each image pixel as belonging to nuclei or surrounding tissue. The result of clustering was a binary image where each pixel was labelled as either belonging or not to the nuclei.

2.2.2. Textural features extraction. By scanning each image with a 5×5 pixel window, the centre pixel was assigned to a five-dimensional textural feature vector comprising:

1. Two features generated from the Mallat non-orthogonal discrete wavelet transform of the energy of the first and second detailed images: The wavelet analysis involved decomposing the original image into a series of low-resolution images by applying a low-pass filter; at the same time, the lost information in the blurred images was captured by convoluting the image with a conjugate high-pass filter producing a series of detailed images. It has been shown that quantitative estimation of detailed images provides a means for encoding textural information [17].

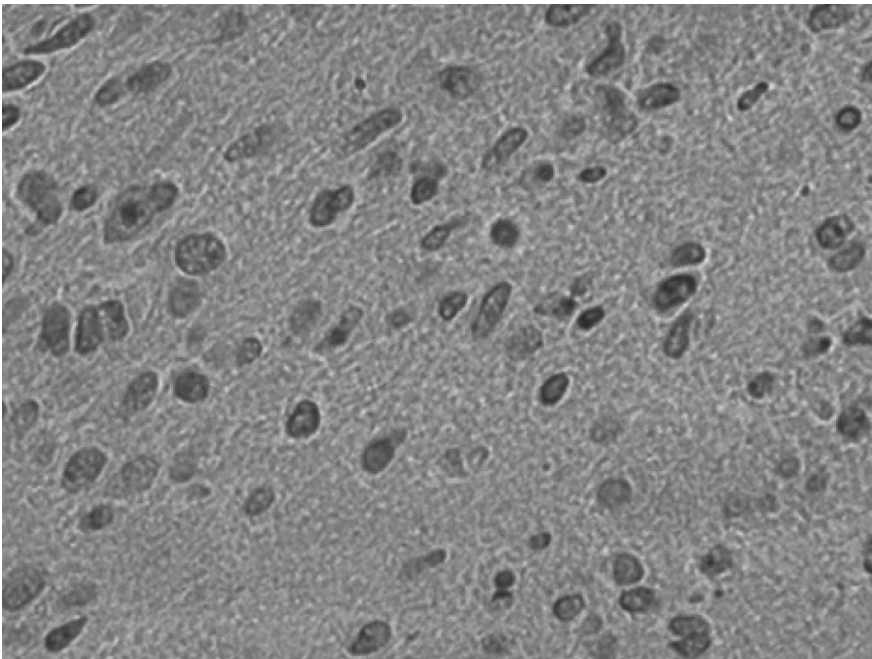


Figure 1. Typical H&E-stained image of astrocytomas.

2. Two features extracted from the autocorrelation function (the spread and cross relation): Because of the inherent characteristic of the autocorrelation function to describe the degree of association between different data points, it has been recognized as a powerful descriptor of chromatin texture [15].
3. One feature motivated by the concept of density estimation using Parzen windows [18]. The parzen estimated descriptor is a new feature we propose for texture estimation that provides an index of pixel inhomogeneities. The mathematical details for calculating these features can be found in Appendix A.

2.2.3. Unsupervised pixel based classification using Support Vector Machine clustering. These textural features involved the input to an unsupervised Support Vector Machine clustering algorithm that classified each image pixel as belonging to the nuclei or background. The basic idea behind SVM clustering, originally proposed in [16], is as follows. The machine initially maps the input data into a higher-dimensional feature space through a nonlinear transformation function. In that feature space, the minimum data enclosing sphere is constructed. By mapping the sphere in the original data space, several contours appear that surround different clusters of data. The Gaussian kernel [19] was utilized as a transformation function. Adjustable parameters were the cost factor C that accounts for outliers and the Gaussian standard deviation (σ). Range values for C and σ tested were $1 \leq C \leq 100$ and $0.05 \leq \sigma \leq 2$. Optimal values were experimentally determined equal to $C = 10$ and $\sigma = 0.5$.

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

After clustering was completed, a binary image was generated where each pixel was labelled as either belonging (white) or not (black) to the nucleus (Figure 2). The resulting binary images were further processed for the elimination of small, noisy regions and corrupted nuclei across image boundaries by means of size filtering and morphological operations [20] (Figure 3). Morphological operations included opening and closing (Equations 2–3):

$$\text{Opening } O(A, B) = A \circ B = D(E(A, B), B) \quad (2)$$

$$\text{Closing } C(A, B) = A \bullet B = E(D(A, -B), -B) \quad (3)$$

where A denotes the original image, B the structuring element, D dilation and E erosion. The final segmented image derives the original image by superimposing with the morphological corrected binary image and performing an *AND* operation (Figure 4).

The best feature vector and the overall performance of the algorithm were defined by evaluating the processed images by the pathologists. By visually inspecting each pair of original and segmenting image, the physicians concluded on the success of the segmentation algorithm in terms of correct and wrongly delineated nuclei. Finally, as an alternative to the SVM clustering algorithm, k-means and fuzzy logic clustering [21] were evaluated (Appendix B).

2.3. Automatic grade classification

2.3.1. Design of the classification scheme. Grade classification scheme was built as a two-level hierarchical decision tree (Figure 5). The first level expressed the system's ability to correctly



Figure 2. Binary image obtained after applying the unsupervised SVM clustering-based pixel segmentation algorithm.



Figure 3. Small and noisy region elimination by size and morphological filtering. The white circle encloses two fused nuclei wrongly recognized as one.



Figure 4. Final segmented nuclei image. The white circle encloses two fused nuclei wrongly recognized as one.

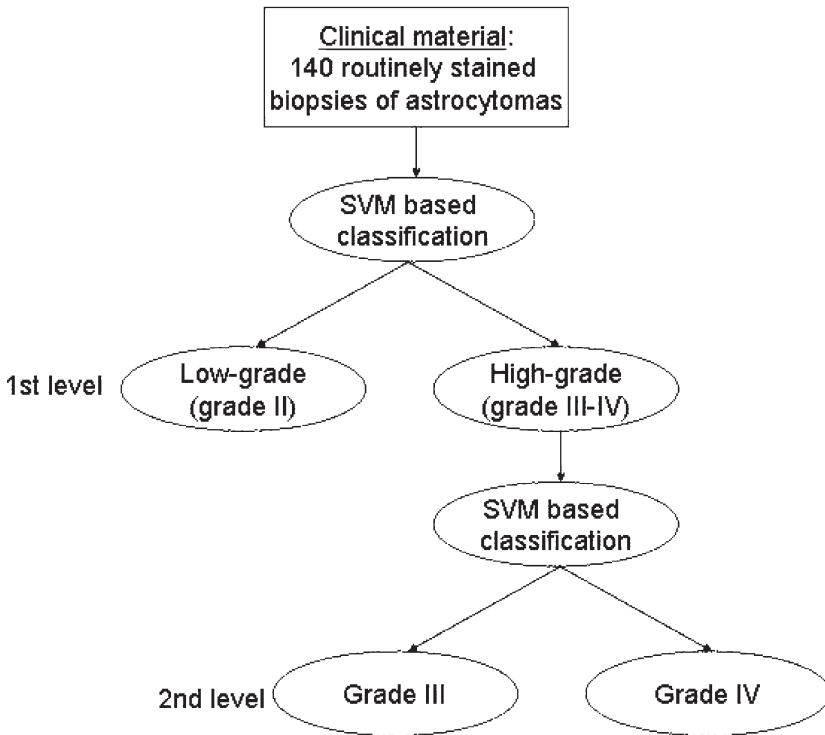


Figure 5. SVM-based decision-tree classification methodology.

classify an unknown case as low-grade (grade II) or high-grade (grade III–IV). At the second level, correctly classified high-grade cases were further categorized as grade III or IV tumours. The decision rule at each level was based on an SVM classification method comprising two distinct stages: (1) feature extraction and (2) feature selection and classification.

2.3.2. Feature extraction. After segmentation, a set of 40 morphological and textural features was extracted describing each individual case (patient). In our previous work, we have shown the importance of these features in encoding information concerning tumours malignancy status [14]. Morphological features describe the shape and size of nuclei (18 features) and comprised area, roundness and concavity. For each of these features, the mean value, standard deviation, range, skewness, kurtosis, and maximum value were calculated. Textural features (first-order, co-occurrence, run-length based) encoded chromatin distribution (22 features) [22]. These 40 dimensional feature vectors were used as input to train and test the classification methodology.

2.3.3. Feature selection and classification. Ninety-two cases (49 GII, 21 GIII and 22 GIV) collected from the University Hospital of Patras were used to construct the SVM classifier at both levels of the decision tree. A detailed description of basic principles behind the SVM classifier and its setting details for the current application can be found in Appendix C.

Exhaustive search was performed in order to determine the best feature vector combination that led to the smallest classification error [23]. Classifier performance was evaluated using the leave-one-out method [21]. Forty-eight cases (12 GII, 9 GIII and 27 GIV) collected from the METAXA hospital were used to validate the systems' generalization to new clinical data. For each level of the decision tree, the best vector combination and the parameters for the SVM classifier that derived from the training stage were retained. The results for both clinical datasets are listed in Tables I and II.

2.3.4. Comparative evaluation of SVM with other machine learning methods. SVM were comparatively evaluated with the Bayesian classifier, Probabilistic neural network (PNN), k -nearest-neighbour classifier, regression and robust estimation. The Bayesian classifier tries to minimize the total classification error assuming normal probability distribution of the classes' samples [5]. On the other hand, the PNN makes no a priori assumption for the underlying probability models of pattern samples. The network estimates each class probability density function (PDF) by linearly combining Gaussian kernel PDF estimation for each training sample for a given class. The standard deviation of the Gaussian determines the sensitivity of

Table I. Training performance of SVM and other machine learning methods in the classification of astrocytomas (training data comprised 92 cases collected from the University hospital of Patras).

Classification task (%)	SVM with RBF kernel (best sigma = 1)	SVM with polynomial kernel of degree d				PNN	Bayesian	k nearest neighbour	Robust estimation	Regression
		$d=1$	$d=2$	$d=3$	$d=4$					
Low-grade	85.7	85.7	89.8	87.7	85.7	87.7	78.7	78.7	77.1	78.7
High-grade	86.7	83.7	90.7	89.8	93.0	81.4	81.0	79.7	67.2	60.8
Overall Accuracy	86.9	84.8	90.2	89.0	89.0	84.8	80.0	79.3	70.4	68.6
Grade III	85.7	76.2	85.7	86.4	80.0	81.0	73.3	73.3	70.0	63.3
Grade IV	86.4	81.8	86.4	90.4	91.8	86.4	80.0	73.3	70.0	70.0
Overall Accuracy	86.1	79.1	86.1	88.3	88.3	83.7	81.0	79.1	77.2	74.7

Table II. Validation performance of the SVM classifier with optimum parameter configuration (validation data comprised 48 cases collected from the METAXA hospital).

Classification task (%)		Optimum SVM configuration: Polynomial kernel of degree $d=2$, $C=10$, support vectors = 16
Low- versus high-grade tumours	Low-grade	85.7
	High-grade	88.8
	Overall accuracy	87.5
Grade III versus grade IV tumours		Optimum SVM configuration: Polynomial kernel of degree $d=3$, $C=10$, support vectors = 10
Grade III versus grade IV tumours	Grade III	82.1
	Grade IV	85.1
	Overall accuracy	83.8

the PNN: too large a standard deviation smoothes out details, whereas too small a standard deviation causes a very spiky approximation representing the k -nearest-neighbour classifier [18]. With the k -nearest-neighbour classifier, the density distribution of the training data is estimated using a distance measure (in this work Mahalanobis [21]). The only adjustable parameter is the number of k neighbours, which also determines the smoothness of the density estimation [6]. Regression analysis was based on the least-square fit method. Robust estimation performs more effectively when data are contaminated with gross errors. An m -type robust estimator based on the Welsch influence function [24] was examined for the current application. Comparative results for all classification methods are recorded in Table I.

3. Results

3.1. H&E image segmentation

The segmentation algorithm was applied to 140 H&E histological images of astrocytomas. The optimum parameter settings for the SVM were $C=10$ and a Gaussian kernel with $\sigma=0.5$. According to the pathologist evaluation, the algorithm correctly recognized and delineated 95% of all nuclei. About 5% of nuclei were erroneously or inadequately circumscribed. Figure 1 illustrates a typical H&E image of astrocytomas. The binary image in Figure 2 consists of nuclei (white) and surrounding tissue (black), as recognized by the SVM clustering algorithm with the best feature vector the energy of the first detailed image of the Mallat wavelet transform, the spread of the autocorrelation function and the parzen estimated descriptor. Fuzzy and k-means clustering were also evaluated for the best feature combination (Appendix B). ‘Damaged’ nuclei located at the image boundaries, small (less than 50 pixels) and noisy regions were omitted by size filtering and morphological operations (Figure 3). In Figure 4, the final segmentation result is illustrated by superimposing the original image (Figure 1) with the final corrected binary image (Figure 3).

3.2. Grade classification

3.2.1. System design. For the first level of the decision tree, SVM optimum configuration was obtained with polynomial kernel of degree 2, $C=10$ and 16 support vectors. The best feature

vector combination comprised three textural features (inertia, inverse different moment and correlation) and two morphological (roundness and the kurtosis of roundness). Low- and high-grade tumours were classified with 89.8 and 90.7% accuracy, respectively. The overall performance was 90.2% (Table I).

Considering the second level of the decision tree, the best SVM configuration was obtained with a polynomial kernel of degree 3, 10 support vectors and $C=10$. This configuration optimized classification performance, resulting in 86.4% accuracy in correctly identifying grade III and 90.4% grade IV tumours. The overall accuracy was 88.3% (Table I). The best feature vector combination consisted of three textural features (inertia, energy and inverse different moment) and one morphological (kurtosis of roundness).

3.2.2. System validation for new clinical data. Retaining the SVM model configuration and using the selected feature vector at each level of the decision tree, the system performance was evaluated by 48 new cases collected from the METAXA hospital. The success rates for low- and high-grade tumour identification were 85.7 and 88.8%, respectively. Grade III tumours were identified with 82.1% success, and grade IV tumours with 85.1% success (Table II).

4. Discussion

In order to account for the subjectivity of the clinicians in grading astrocytomas, several computer-based studies have been proposed. However, most of these studies have been based on specialized clinical protocols [5–11]. One of the reasons has been that H&E-stained images are difficult to segment [15,16]. To support the regular diagnostic procedure followed by the experts, we investigated and developed an automatic segmentation and classification methodology compatible with every-day clinical standards, namely the WHO grading and H&E-staining protocol.

One hundred and forty H&E-stained images of astrocytomas were segmented, employing an unsupervised SVM clustering methodology. According to the pathologists' evaluation, nuclei were accurately circumscribed, retaining their morphology and boundary characteristics. This is mainly due to the fact that the algorithm took into account the subtle differences in chromatin grey-level distribution in nuclei and surrounding tissue. These subtle differences in the grey-level allocation were encoded in the form of five textural features based on the wavelet Mallat transform, the autocorrelation function and a parzen estimated feature. Since the algorithm considered solely local information, it was insensitive to biopsy preparation parameters, such as variations in the amount of stain concentration in each individual image or even within different regions in the same image. The small white pixel clumps within the background (Figure 2) appear because not only does the H&E complex target nuclei but it also concentrates in a variety of structures outside nuclei; this targeting behaviour complicates image interpretation without, however, providing any additional diagnostic information. The elimination of these problematic regions was performed by morphological operations (Figure 3). As can be seen in Figure 4, fused or overlapped nuclei were incorrectly recognized as one structure by the algorithm. Nevertheless, the recognition of such structures is of minor significance, since even during visual inspection, pathologists omit these nuclei because they are difficult to interpret. The proposed algorithm was validated by two pathologists by comparing the segmented images with the original images and marking wrongly segmented nuclei. In this perspective, 95% of all delineated nuclei on the tested images were circumscribed correctly and 5% incorrectly. Considering that segmented nuclei from each section field ranged

between 45 and 95, the misclassification error of 5% may be regarded as of limited significance.

As far as the best feature combination that optimized the performance of the SVM clustering algorithm, it comprised the energy of the first detailed image of the Mallat wavelet transform, the spread of the autocorrelation function and the parzen estimated feature. The first two of these features have already been recognized as powerful descriptors for encoding texture. The parzen estimated descriptor is a new feature we propose for texture estimation; it provides an index of pixel inhomogeneities. If the pixels in the examining window have similar intensity values, then the parzen estimation is high. When inhomogeneities are detected (i.e. boundaries), then the pixel values vary significantly, giving lower values for the parzen descriptor (Appendix A).

The proposed segmentation algorithm comprises the continuation of our work in H&E-stained image segmentation. In [14], we have examined the effectiveness of a supervised pixel-based classification algorithm evaluating features extracted from the autocorrelation function for the segmentation of H&E images. The algorithm has been based on training a neural network classifier to identify nuclei and surrounding tissue, based on the autocorrelation differences of randomly and manually sampled nuclear and background image regions. In the current approach, we extended our research to unsupervised methods and examined a variety of textural features. Previous work in H&E-stained image segmentation is relative limited; the intense diversity of this staining process in targeting nuclei makes the segmentation task difficult [14,25]. Among most recent approaches, Mouroutis *et al.* [25] proposed a technique for H&E image segmentation using statistical modelling. The algorithm was based on the Hough transform for estimating possible nuclei locations and the likelihood function for optimizing nuclear delineation. An initial hypothesis had to be established assuming a priori knowledge of the morphological characteristics of cell nuclei, which derived from a visual inspection of each image separately. Furthermore, the algorithm presented some difficulties in recognizing nuclei boundaries when sub-nuclear structures also existed (e.g. nucleoli). These sub-nuclear structures were initially compensated by Gaussian filtering. In our approach, we make no a priori assumptions concerning nuclear morphology, texture or the presence of smaller structures within nuclei; the whole procedure runs completely with the need for supervision. To the best of our knowledge, this is the first time texture-based unsupervised segmentation of H&E-stained images has been examined.

The state-of-the-art SVM-clustering algorithm [16] accurately identified the different pixel populations of nuclei and surrounding tissue. The advantages of this algorithm are: (1) no a priori assumption for the number of clusters is needed, as the algorithm arbitrary forms several clusters without any kind of bias; (2) by carefully adjusting the cost factor, C , the SVM can account for the existence of outliers (e.g. regions with very intense or inadequate staining (Appendix B)); (3) clustering boundaries are of an arbitrary shape in contrast to classical clustering algorithms that form geometrically defined boundaries (e.g. ellipsoid, etc.).

After segmentation, a 40-dimensional feature vector was extracted for each case and evaluated by an SVM decision-tree classification scheme. The decision-tree structure was preferred to using a single three-class SVM classifier because it resembles the diagnostic procedure followed by the pathologist, and it is less complex and faster to implement. Low-grade tumours were distinguished from high-grade tumours with an accuracy of 90.2% and grade III from grade IV with 88.3%. Although data were collected from two different hospitals, and considering that the staining protocol utilized in this study is not as accurate as the specialized techniques investigated in previous studies [5–11], the results

are most encouraging and in any case comparable with those recorded in the literature. Schad *et al.* [5] evaluated Feulgen-stained microscopic slides with 75% success in correctly identifying astrocytomas grade. Using the same staining process, Decaestecker *et al.* [6] proposed a nearest-neighbour classification technique with 55% success rate. Belacel *et al.* [9] presented a fuzzy-logic system analysing H&E-stained images with 66% discrimination accuracy. Kolles *et al.* [10] proposed a new specialized grading system (HOM system) evaluating the combination of Feulgen and ki-67-stained biopsies. Although using the HOM system classification accuracy was $>90\%$, when the authors tested their system with the WHO scheme, performance was reduced to about 60%. Among the most recent approaches, Reinhold *et al.* [7] analysed ki-67-stained images and proposed an automatic classification system that distinguished only low- (grade II) from high-grade (grade III) tumours with 88% accuracy. Our method combines relative high accuracy and every-day clinical protocols. Thus, it may be utilized in daily clinical practice without the need to perform special staining techniques (that moreover complicate and increase the time/cost of diagnosis) or the need for pathologists to familiarize with new grading protocols.

Furthermore, the SVM classifier was comparatively evaluated with several machine-learning methods. The superiority in performance of the SVM for both levels of the decision tree (Tables I and II) might be explained as follows: (1) it does not assume any a priori knowledge of the underlying probability functions; (2) it does not risk becoming trapped in local minima; (3) it can account for gross errors that do not fit the stochastic model of nuclei and background pixel allocations (outliers). On the other hand, SVM optimization is complicated and time-consuming, and its setting is much more elaborate compared with the rest of techniques examined in this work. Nevertheless, the machine generalized well on new clinical data; this can be verified both from the high classification accuracy on the new clinical dataset and from the relative small numbers of support vectors identified for each level of the decision tree. For the unseen dataset, low-grade cases were distinguished from high-grade cases with 87.5% accuracy with 16 support vectors (11.4% of all cases). Grade III were distinguished from grade IV tumours with 83.8% accuracy with 10 support vectors (12.7% of all high-grade cases)

As far as the clinical problem is concerned, most previous studies have presented techniques that have mostly focused exclusively in distinguishing low- from high-grade tumours [5–8]. This was mainly because, until recently, the belief has been that grade III and IV tumours must be treated identically [1]. However, new studies have revealed that it is most important to accurately define groups of patients with high-grade astrocytomas, because grade III tumours are more chemosensitive than grade IV [12–13]. Responding to this advancement, our algorithm was designed to perform well in distinguishing low- from high-grade tumours, and grade III from grade IV tumours. Under this perspective, although the polynomial kernel of degree 4 optimizes overall classification performance in discriminating grade III from grade IV tumours, it is not selected as the optimum because of its reduced performance in identifying grade III tumours (Table I). On the other hand, the polynomial kernel of degree 3 resulted in high classification performance for both levels of the decision tree.

To summarize, in this work we have developed a computer-based method for the automatic characterization of the degree of tumour malignancy for brain astrocytomas. Segmentation and classification results are very encouraging, considering that the method was developed based on every-day clinical standards. The proposed methodology might be used in parallel with conventional grading to support the regular diagnostic procedure and reduce subjectivity in astrocytomas grading.

References

1. Lisa M. Brain tumours. *The New England Journal of Medicine* 2001;344:114–123
2. World Health Organization. International histological classification of tumours: Histological typing of tumours of the central nervous system, 2nd ed. Berlin: Springer, 1993
3. Prayson RA, Agamanolis DP, Cohen ML, Estes ML. Interobserver reproducibility among neuropathologists and surgical pathologists in fibrillary astrocytoma grading. *Journal of the Neurological Sciences* 2000;175:33–39.
4. Coons W, Jhonson P, Sceithauer B, Yates A, Pearl D. Improving diagnostic accuracy and interobserver concordance in the classification and grading of primary gliomas. *Cancer* 1997;79:1381–1393.
5. Schad LR, Schmitt HP, Oberwittler C, Lorenz WJ. Numerical grading of astrocytomas. *Medical Informatics* 1987;12:11–22.
6. Decaestecker C, Salmon I, Dewitte O, Camby I, Van Ham P, Pasteels J, Brotchi J, Kiss R. Nearest-neighbor classification for identification of aggressive versus nonaggressive astrocytic tumours by means of image cytometry-generated variables. *Journal of Neurosurgery* 1997;86:532–537.
7. Reinhold N, Schlote W. Topometric analysis of diffuse astrocytomas. *Analytical and Quantitative Cytology and Histopathology* 2003;25:12–18.
8. Scarpelli M, Montironi R, Thompson D, Bartels P. Computer-assisted discrimination of glioblastomas. *Analytical and Quantitative Cytology and Histopathology* 1997;19:369–375.
9. Belacel N, Boulassel M. Multicriteria fuzzy assignment method: a useful tool to assist medical diagnosis. *Artificial Intelligence in Medicine* 2001;21:201–207.
10. Kolles H, v. Wangenheim A, Rahmel J, Niedermayer I, Feiden W. Data-driven approaches for decision making in automated tumour grading. An example in astrocytoma grading. *Analytical and Quantitative Cytology and Histopathology* 1996;18:298–304.
11. Sallinen P, Sallinen S, Helen T, Rantala I, Rautiainen E, Helin H, Kalimo H, Haapsalo H. Grading of diffusely infiltrating astrocytomas by quantitative histopathology, cell proliferation and image cytometric DNA analysis. *Neuropathology and Applied Neurobiology* 2000;26:319–331.
12. Larner JM. Radiation Therapy Oncology Group. A phase i/iii randomized study of radiation therapy and temozolomide versus radiation therapy and bcnu for anaplastic astrocytomas RTOG 98-13, 2002.
13. Iwadata Y, Fujimoto S, Yamaura A. Differential chemosensitivity in human intracerebral gliomas measured by flow cytometric DNA analysis. *International Journal of Molecular Medicine* 2002;10:187–192.
14. 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. *Medical Informatics And The Internet In Medicine* 2002;27:111–122.
15. Spyridonos P, Cavouras D, Nikiforidis G. Neural Network based segmentation and classification system for the automatic grading of histological sections of urinary bladder carcinoma. *Analytical and Quantitative Cytology and Histology* 2002;26:317–324.
16. Asa Ben-Hur, Horn D, Siegelmann HT, Vapnik V. Support vector machine clustering. *Journal of Machine Learning Research* 2001;2:125–137.
17. Van de Wouwer G, Weyn B, Scheunders P, Jacob W, Van Marck E. and Van Dyck D, Wavelets as chromatin texture descriptors for the automated identification of neoplastic nuclei. *Journal of Microscopy* 2000;97:25–35
18. Specht D. Probabilistic neural networks. *Neural Networks* 1990;3:109–118.
19. Kechman V. Support vector machines. In: *Learning and Soft Computing*. Cambridge, MA: MIT Press; 2001. p. 121–184.
20. Spyridonos P, Ravazoula P, Cavouras D, Berberidis K, Nikiforidis G. Computer-based grading of haematoxylin-eosin stained tissue sections of urinary bladder carcinomas. *Medical Informatics and the Internet in Medicine* 2001;26:179–190.
21. Theodoridis S, Koutroubas K. Feature generation II. In: *Pattern Recognition*. New York: Academic Press; 1999. p. 233–270.
22. Glotsos D, Spyridonos P, Petalas P, Cavouras D, Nikiforidis G. A hierarchical decision tree classification scheme for brain tumour astrocytoma grading using Support Vector Machines. In: *Third International Symposium on Image and Signal Processing and Analysis*. In press.
23. Fellow J, Duijn R, Mao J. Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2000;22:4–37.
24. Stewart CV. Robust parameter estimation in computer vision. *SIAM Review*;41:513–537.
25. Mouroutis, Roberts SJ, Bharath AA. Robust cell nuclei segmentation using statistical modeling. *Bioimaging* 1998;6:79–91.
26. Kalatzis I, Pappas D, Piliouras N, Cavouras D. Support vector machines based analysis of brain SPECT images for determining cerebral abnormalities in asymptomatic diabetic patients. *Medical Informatics And The Internet In Medicine* 2003;28:221–230.

27. Christianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge: Cambridge University Press; 2000.
28. Mathworks. Optimization toolbox, <http://www.mathworks.com/products /optimization/?BB = 1>

Appendix A

Based on the Mallat non-orthogonal discrete wavelet transform [17], each image window was decomposed into a series of low-resolution images by applying a low-pass filter. The information loss resulting from the blurring operation was retrieved by convoluting the original image window with conjugate high-pass filters. This operation yielded detailed images containing details lost by the low-pass filtering. These details were encoded in the form of quantitative measurements; the energy of the first and second detailed image (Equations A1 and A2) has been shown to be a powerful descriptor of texture representation. The energy was calculated as the mean of the squared wavelet coefficients for every detailed image [17].

$$D_s(x,y) = I \frac{\partial \theta(x,y)}{\partial x} \tag{A1}$$

$$W_i = \frac{1}{MN} \sum_{x,y} (D_s(x,y))^2 \tag{A2}$$

where D_s is the wavelet transform of the 5×5 image window $I(x,y)$, θ a smoothing function, $1 < M < x$, $1 < N < y$ = image rows and columns, W_i is the energy of the i th detailed image, and d is the wavelet depth representation.

The features based on the autocorrelation function were the cross-relation $S(1,1)$ and the second-degree spread $S(2,2)$ (Equations A3–A6).

$$S(u,v) = \sum_{m=0}^5 \sum_{n=-5}^5 (m - n_m)^u (n - n_n)^v A_F(m,n) \tag{A3}$$

where

$$n_m = \sum_{m=0}^5 \sum_{n=-5}^5 m A_F(m,n) \tag{A4}$$

$$n_n = \sum_{m=0}^5 \sum_{n=-5}^5 n A_F(m,n) \tag{A5}$$

and

$$A_F(m,n) = \sum_j \sum_k F(j,k) F(j - m, k - n) \tag{A6}$$

is the autocorrelation function for displacements m, n . These features have been suggested for texture estimation, since the autocorrelation function presents the inherent characteristic of encoding the degree of association between different data points [15].

The parzen based estimated feature, accounting for the local pixel inhomogeneities, was computed as follows: (1) the Euclidean distance of the centre pixel from the rest of the pixels

was calculated at each position of the scanning window; (2) the distance vector passed through a Radial basis function kernel; and (3) finally the centre pixel of the window was assigned a feature value deriving by adding the exponentiated distances for all pixels (Equation A7).

$$\sum_{i=1}^{N_k-1} \exp\left(-\left(\frac{d0.8326}{\sigma}\right)^2\right) \quad (\text{A7})$$

where d =Euclidean distance, N =total number of window pixels, σ =the Gaussian standard deviation. The optimum setting for σ was experimentally determined equal to 0.6 after testing values ranging from 0.05 to 2.

Appendix B

Figure B1 depicts the segmentation result of the proposed methodology using alternatively fuzzy, k-means and SVM clustering. SVM clustering performed better in nuclei segmentation because of: (1) more accurate separation of closely located nuclei; (2) elimination of small noisy regions within the background and nuclei; (3) ‘filtering’ of small sized structures which were recognized as outliers; (4) insensitivity to the existence or not of nucleoli or other

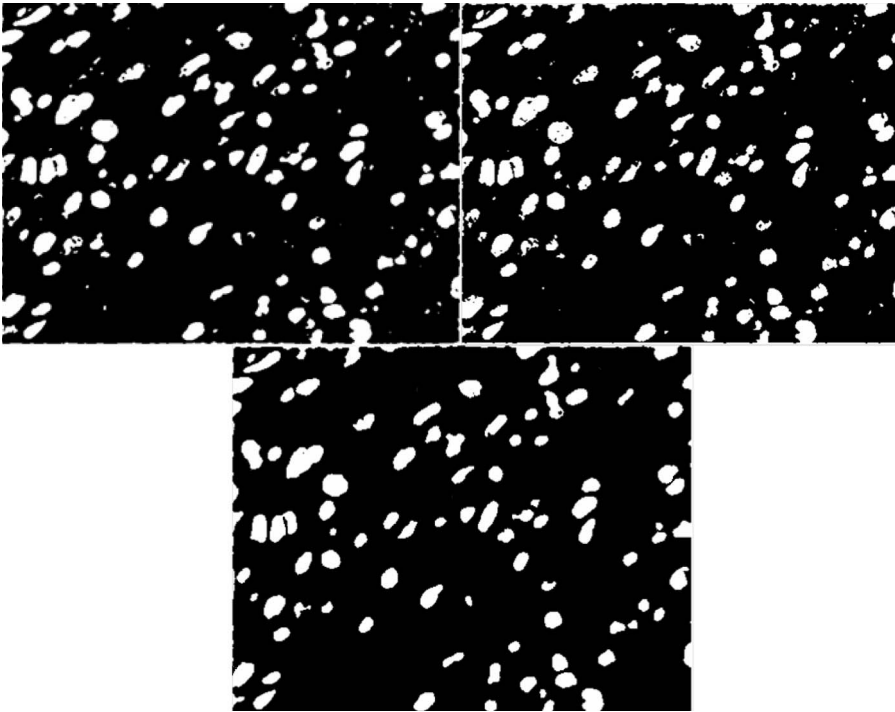


Figure B1. Top: Binary images obtained after applying the proposed segmentation algorithm with fuzzy (left) and k -means (right) clustering. Bottom: Binary image obtained after applying the proposed segmentation algorithm with SVM clustering constructed with RBF kernel ($\sigma = 0.5$). The best feature of vector combination was the energy of the first detailed image of the Mallat transform, the spread of the autocorrelation function and the parzen estimated feature.

sub-nuclei structures; and (5) more accurate nuclei detection according to the pathologists' evaluation compared with the two other algorithms tested.

Appendix C

The basic idea in applying SVM [19,26] in classification problems is twofold: map the input space in a one-feature space (usually of higher dimension) through a linear or nonlinear transformation function (kernel). In this feature space, seek the optimal linear separating hyperplane with maximum distance from the closest training data. These training data are called support vectors. The fewer the number of support vectors, the larger the probability that the classifier might generalize well on unseen data. The decision function of SVM for binary classification problems has the form:

$$g(x) = \text{sign} \left(\sum_{i=1}^N \alpha_i y_i K(x, x_i) + b \right) \quad (\text{C1})$$

where x_i denotes the training data belonging to either class $y_i \in \{+1, -1\}$, N is the number of training samples, α_i , b are weight coefficients, and $K(x, x_i)$ is the kernel function. The most common and admissible kernel functions are the Gaussian RBF and polynomial (Equations C2 and C3).

$$K_{\text{RBF}}(x, x_i) = \exp \left(\frac{-\|x-x_i\|^2}{2\sigma^2} \right) \quad (\text{C2})$$

$$K_{\text{POLYNOMIAL}}(x, x_i) = ((x^T x_i) + 1)^d, \quad d = \text{degree} \quad (\text{C3})$$

The determination of the maximal margin hyperplane and weight parameters is a complex nonlinear quadratic optimization problem with inequality constrains. In the current application, the optimization problem was solved by using the routine quadprog [27] provided with the MATLAB optimization toolbox [28]. SVM were constructed with the RBF and polynomial kernels of degree 1, 2, 3, 4 as mapping functions. Concerning the Gaussian standard deviation, values ranging from 0.05 to 2 were tested. The optimum value of sigma was found to be equal to 1. Finally, the cost factor, C , which determines the penalty for misclassifications, was experimentally set equal to 10.

