

AUTOMATED DIAGNOSIS OF BRAIN TUMOURS ASTROCYTOMAS USING PROBABILISTIC NEURAL NETWORK CLUSTERING AND SUPPORT VECTOR MACHINES

DIMITRIS GLOTSOS

*Department of Medical Physics, University of Patras,
 Rio-Patras, 26500, Greece
 dimglo@med.upatras.gr*

JUSSI TOHKA

*Tampere University of Technology, Institute of Signal Processing, Tampere, Finland
 and
 University of California Los Angeles, Department of Neurology,
 Laboratory of NeuroImaging, USA
 jussi.tohka@tut.fi*

PANAGIOTA RAVAZOULA

*University Hospital of Patras, Department of Pathology,
 Rio-Patras, 26500, Greece*

DIONISIS CAVOURAS

*Technological Institute of Athens, Department of Medical Instruments Technology,
 Aigaleo, Athens, 12210, Greece
 cavouras@teiath.gr*

GEORGE NIKIFORIDIS

*Department of Medical Physics, University of Patras,
 Rio-Patras, 26500, Greece
 gnikif@med.upatras.gr*

A computer-aided diagnosis system was developed for assisting brain astrocytomas malignancy grading. Microscopy images from 140 astrocytic biopsies were digitized and cell nuclei were automatically segmented using a Probabilistic Neural Network pixel-based clustering algorithm. A decision tree classification scheme was constructed to discriminate low, intermediate and high-grade tumours by analyzing nuclear features extracted from segmented nuclei with a Support Vector Machine classifier. Nuclei were segmented with an average accuracy of 86.5%. Low, intermediate, and high-grade tumours were identified with 95%, 88.3%, and 91% accuracies respectively. The proposed algorithm could be used as a second opinion tool for the histopathologists.

Keywords: Probabilistic neural network; support vector machines; microscopy; astrocytomas; grading.

1. Introduction

Astrocytomas are considered to be among the most lethal and difficult-to-treat forms of cancer.¹ In diagnosing of astrocytomas, the most significant step is

the determination of the degree of tumour abnormality (grading). Grading is performed by histopathologists who visual inspect microscopic sections of biopsy material under the microscope.² According to the World Health Organization (WHO) grading

system,³ astrocytomas are classified into three grades (grade II–III and IV). Astrocytomas of grade II are considered as low-grade tumours, which generally have good prognosis. Astrocytomas of grade III and IV (high grade) are the most aggressive tumours with survival time ranging on average from 6 to 12 months.⁴

Although the WHO grading system is the most popular for grading astrocytomas, the WHO guidelines are limited by the vagueness of the descriptions used to define each grade. Some of these definitions become clearer only by experience.⁵ The latter promotes relatively low inter and intra-observer reproducibility among histopathologists, especially in the critical separation of grade II from grade III tumours, because malignancy from grade II to grade III tumours develops along a biological continuum.⁶ From a clinical point of view it is important to distinguish between low and high-grade tumours, since many high grade astrocytomas are often treated differently (frequently with radiation therapy) compared to low-grade astrocytomas.⁴ Hence, there is a need for the development of reliable and reproducible systems for assessing tumor aggressiveness.

Although most studies have focused on validation of visible features,⁷ a variety of automatic techniques have been introduced to improve diagnostic accuracy.^{8–12} Reinhold *et al.*⁸ have evaluated with discriminant analysis the changes in topometry of nuclei and developed a system that discriminated grade II from grade III tumours with an accuracy of 88%. Decaestecker *et al.*⁹ used a nearest-neighbor classifier that have discriminated less aggressive from highly aggressive low-grade astrocytomas with 55% accuracy. Scarpelli *et al.*¹⁰ showed a significant change in quantitative nuclear features from low to high-grade tumours using linear discriminant analysis, toluidine-blue-staining and the Burges grading system. Wangenheim *et al.*¹¹ have developed a new grading system (named HOM) that has been shown to be more efficient in automatic grading of gliomas compared to the WHO grading system. In Wangenheim’s study, different grade gliomas were discriminated with 97% accuracy using the HOM system, with 49% accuracy using the Daumas-Duport system, and with 61% accuracy using the WHO grading system. Belacel *et al.*¹² have developed a fuzzy logic method for astrocytomas grading

that automatically classified different tumour grades with an accuracy of 66%.

However, many of these studies have introduced supervised computer-based systems^{8–12} that were developed based on modifications of the WHO grading scheme^{10–11} and specialized staining protocols.^{8–11} Additionally, these studies have not investigated the possibility of ‘suspicious’ cases, which are those critical cases that are characterized by histopathologists as ranging from grade II (low grade) to grade III (high grade). It is most significant to accurately specify these ‘intermediate’ grade tumours (between grade II and grade III), because these tumours need re-examination. In this study we propose a novel method designed for the automatic segmentation and classification of astrocytomas microscopy images to be compatible to the most widely accepted clinical protocols for astrocytomas grading, namely the WHO scheme and the Hematoxylin-Eosin (H&E) staining protocol.^{3,13} In addition, the proposed method (i) is designed in an unsupervised manner to perform image segmentation by using a novel Probabilistic Neural Network (PNN) pixel-clustering algorithm (ii) uses quantitative nuclear features and a Support Vector Machine (SVM) classifier for the automatic separation of low from high grade tumours, and (iii) identifies those intermediate-grade tumours that need re-examination using a decision tree classification scheme.

2. Material and Methods

One hundred and forty (140) H&E stained biopsies of astrocytomas were collected from the Department of Pathology of the University Hospital of Patras, Greece. Tumour grade was defined as low (61/140), high grade (67/140) or as ‘suspicious’ cases of intermediate grade (12/140) (see Table 1) according the WHO grading system by a histopathologist (P.R.).

Table 1. Astrocytomas biopsies dataset.

Low grade	Suspicious cases of intermediate grade	High grade	
Grade II	Grade II–III	Grade III	Grade IV
61	12	29	38

For each biopsy, the histopathologist specified the most representative region. From this region, images (Fig. 4) were acquired ($768 \times 576 \times 8$ bit) using a light Zeiss Axiostar-Plus microscope (ZEISS; Germany) connected to a Leica DC 300 F color video camera (LEICA; Germany).

2.1. Image segmentation

Nuclei were automatically segmented from surrounding tissue in order to encode tumour malignancy by means of a set of features derived from the segmented nuclei. To design the segmentation algorithm 5×5 windows around each pixel were considered. The center pixel of each window was assigned a 3-dimensional textural feature vector based on which pixels were classified as belonging to nuclei or surrounding tissue. The textural features were (1) the sum of autocorrelation function A_F for all possible displacements m, n inside the window 5×5 , (2) the cross relation $S(1,1)$ and (3) the second degree spread $S(2,2)$.¹⁴

$$S(1,1) = \sum_{m=0}^5 \sum_{n=-5}^5 (m - n_m)(n - n_n) A_F(m, n) \quad (1)$$

$$S(2,2) = \sum_{m=0}^5 \sum_{n=-5}^5 (m - n_m)^2 (n - n_n)^2 A_F(m, n) \quad (2)$$

$$A_F(m, n) = \sum_j \sum_{k'} F(i, j) F(i - n, k - m) \quad (3)$$

where $n_m = \sum_{m=0}^5 \sum_{n=-5}^5 m A_F(m, n)$, $n_n = \sum_{m=0}^5 \sum_{n=-5}^5 n A_F(m, n)$, and $F(i, j)$ is the image intensity at the pixel (i, k) .

From Eq. (3) it is apparent that a window region will exhibit higher sum of the autocorrelation function for a fixed shift (m, n) than a region of fine texture. The sum of autocorrelation function for all displacements over the window 5×5 could be regarded as coarseness index; this feature takes higher values for coarser regions. Since the sum of autocorrelation function is estimated for all displacements, it would be possible that two different textural fields will exhibit the same value. To improve the texture discrimination, the textural features $S(1,1)$ and $S(2,2)$ were estimated additionally, which provide the extra information of two-dimensional spread

measures of the autocorrelation function.¹⁴ In this way formation of clusters (nuclei and surrounding tissue) was facilitated.

The feature vectors (sum of the autocorrelation function, spread and cross relation) generated for the window around each pixel were fed into a Probabilistic Neural Network (PNN)-based clustering algorithm¹⁵ (see Sec. 2.1.1.). The algorithm assigned the central pixel of each window as belonging to either nuclei or surrounding tissue based on the feature vector of that window. In the resulted binary image (nucleus and surrounding tissue), pixels that were classified as emanating from surrounding tissue were marked with black colour and those belonging to a nucleus were marked with white colour (Fig. 5). To reduce noise the binary image was further processed by fill holes filters, morphological operations, elimination of nuclei less than 200 pixels and elimination of nuclei intersecting image boundaries (Fig. 6).¹³ Finally, nuclei segmentation was achieved by superimposing the binary processed image to the original image (Fig. 7).

To evaluate the performance of the segmentation procedure, nuclei boundaries in 25 randomly selected images were manually delineated by an experienced histopathologist (P.R.). Subsequently, manually segmented nuclei were compared to the automatically segmented nuclei in terms of nuclei area, roundness and concavity.

2.1.1. Probabilistic Neural Network-based robust clustering

In order to implement PNNs for robust clustering, the network architecture is constructed with 5 (input layer, pattern layer, summation layer, clustering layer and output layer) instead of 4 layers as originally proposed in Ref. 16 (Fig. 1).

The *input layer* comprises as many nodes as the number of normalized to zero mean and unit deviation¹⁷ input feature (or pattern) vectors. The *pattern layer* consists of as many nodes as the input feature vectors; each feature vector corresponds to exactly one pattern node. At each pattern node, the Euclidean distances of the node's corresponding feature vector from the rest of the input feature vectors are calculated. The resulting Euclidean distances vector is then passed through a Gaussian

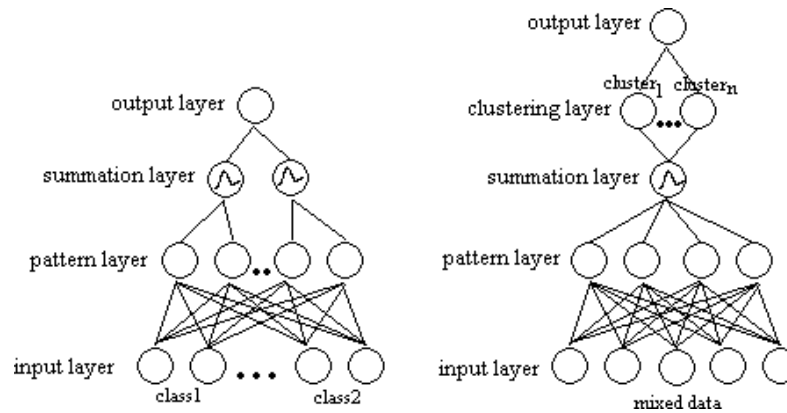


Fig. 1. PNN architecture for supervised (left) and unsupervised (right) classification.

activation function and summed to give an estimation of the PDF $f(x_i)$ at the location of each input feature vector. In other words, at each node i of the pattern layer we calculate

$$f(x_i) = \sum_{j=1}^N \exp\left(-\frac{\|d\|^2}{2\sigma^2}\right) \quad (4)$$

where $x_i = S(1,1) + S(2,2) + \text{sum}(A_F)$, d is the Euclidean distances vector and σ is the spread. In the *summation layer* the whole data PDF is calculated by combining the outputs from the pattern layer and interpolating. In symbols, the PDF is estimated based on the individual PDF estimates f_i :

$$f(y; x_1, \dots, x_N). \quad (5)$$

The value of the $f(y_i) = f_i$ and the values other than the given input feature vectors are obtained by linear interpolation. The spread is an adjustable parameter, which has been shown¹⁸ to play an important role in the estimation of the data PDF in Eq. (4). Too small σ causes very spiky approximation, whereas too large smoothes out details. Although the optimal value of σ is usually experimentally determined, in this work we estimated σ according to:

$$\sigma = \frac{1}{N-1} \sum_{i=1}^N \sum_{j=1}^N \|x_i - x_j\|, \quad i \neq j. \quad (6)$$

In this way sigma is independent of the type of data and additionally sigma is automatically estimated.

In the *clustering layer* (Fig. 1b), the different clusters are identified. Initially, peaks corresponding to class (cluster) centroids c_i are identified on the 1-dimensional representation of the PDF (see Fig. 2).

The most prominent peaks of the PDF are first determined by locating the most prominent peak that corresponds to the first cluster centroid. This peak and its surrounding points are then removed from consideration. The surrounding points are defined within a Gaussian estimated spread calculated as:

$$\text{spread} = \frac{1}{M} \max(\|x\|)^2 \quad (7)$$

where M is the number of data clusters. The algorithm then continues by specifying and removing the second peak, that corresponds to the second cluster centroid and this iterative procedure continues by locating cluster centroids (c_1, \dots, c_M) as many times as the *a priori* estimated number of data clusters M . This implies that we have a vague idea of the classes we are seeking for. A schematically view of this procedure is illustrated in Fig. 2. Clusters are then, formed by assigning each data point x_i to cluster c_i if the majority of its K nearest neighbors is closer to c_i than any other cluster centroid. K is defined as $K = 2 + \frac{d_{\max}}{\sqrt{2M}}$, where d_{\max} is maximum Euclidean distance between cluster centroids. Before assigning all points to clusters, a procedure to detect outlying points is initiated. More distant probable outliers are considered those feature vectors located at the two ends of the PDF satisfying $\max(\|x\|)^2$ (far right O_R and far left O_L edges in Fig. 2d). Feature vectors that are closer to O_R or O_L than c_1 or c_2 in Fig. 2d are considered outliers, provided that these feature vectors are outside the corresponding to c_1 or c_2 spreads (see Eq. (7)). These outliers are considered as belonging to no class. Finally, in the *output layer* each data sample is assigned as belonging to cluster c_j or being an outlier.

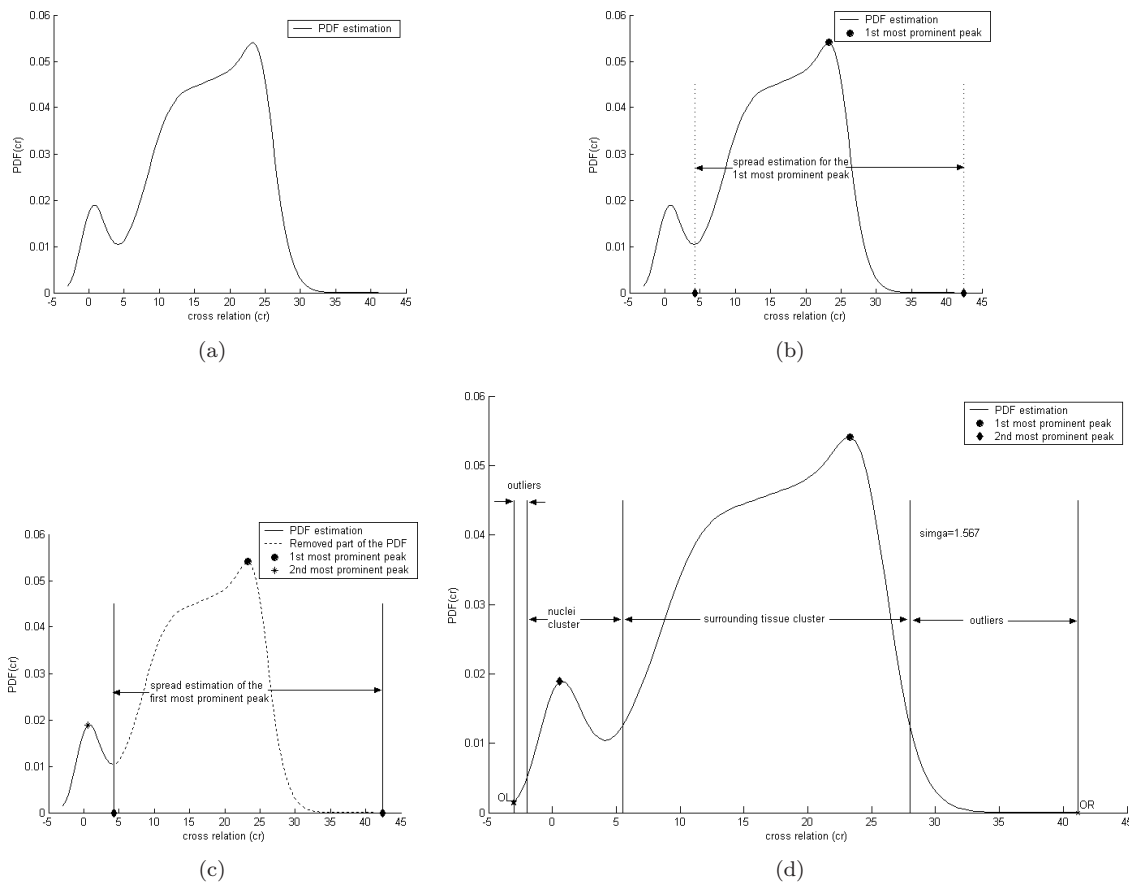


Fig. 2. (a) Calculation of data PDF, (b) Identifying the first most prominent peak, (c) Removing this peak and its surrounding points within a Gaussian estimated spread according to Eq. (7) and calculating the second most prominent peak, (d) Formation of clusters and identification of outliers using a k -nearest neighbor procedure.

2.2. Image classification

After image segmentation, 33 morphological and textural features (Appendix A) were extracted for each case (patient) by selecting at least 50 non-overlapping nuclei. Textural features (15 features) described chromatin organization inside nuclei.^{19–20} Morphological features (18 features) encoded nuclei size and shape.²¹ A two level hierarchical decision tree (Fig. 3) was constructed for the automatic identification of initially low from high-grade tumours and subsequently suspicious cases of intermediate grade in the low and high-grade groups respectively. At each level of the decision tree an SVM classifier²² was utilized alternatively constructed with an RBF and polynomial kernels. To reduce the dimensionality of the features an exhaustive search algorithm was performed in all possible combinations up to 6. The performance of the SVM classifier for each feature

combination was assessed by using a leave-one-out method.²³

2.2.1. Support Vector Machine classification

By mapping input vectors into a higher dimension feature space and defining the hyperplane that has the maximum distance from the closest training data, SVM can be utilized for binary classification problems^{24–25} with discriminant function as follows:

$$g(x) = \text{sign} \left(\sum_{i=1}^N \alpha_i y_i K(x, x_i) + b \right) \quad (8)$$

where x_i training data belonging to either class $y_i \in \{+1, -1\}$, N the number of training samples, α_i , b weight coefficients and K the transformation or kernel function. Kernel functions utilized were the radial basis function (RBF) with value of $\gamma = 1/(2\sigma^2)$ set equal to 0.5 after testing values from 0.005 to 6 and

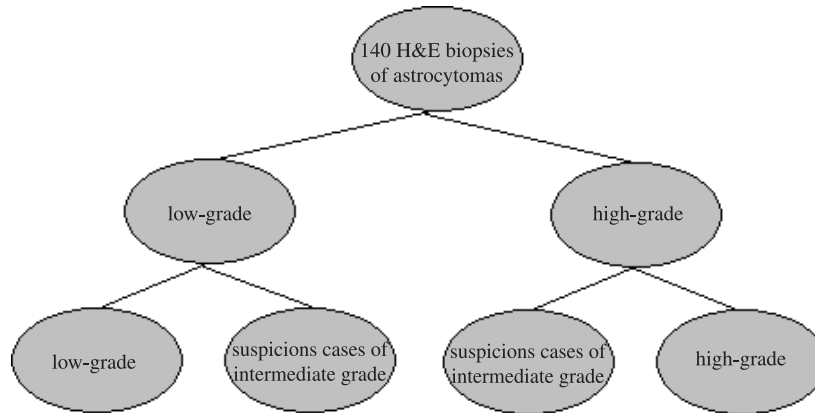
6 *D. Glotsos et al.*

Fig. 3. Decision tree classification scheme for discriminating low, high and intermediate cases of astrocytomas.

polynomial of degree $d = 1$ and 2. The adjustable parameter C that specifies the importance of misclassifications was experimentally determined equal to 10.

$$K_{\text{REF}}(x, x_i) = \exp\left(\frac{-\|x - x_i\|^2}{2\sigma^2}\right), \quad (9)$$

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

Using the routine quadprog provided with the MATLAB optimization toolbox, optimization problem of calculating parameters a_i was solved.

3. Results

Figure 4 illustrates a digitized H&E stained image ($768 \times 576 \times 8$ bit) of a low-grade astrocytic tumor. The pixel classification procedure performed by the PNN algorithm resulted in the binary image of Fig. 5. In Fig. 6 nuclei were corrected by applying

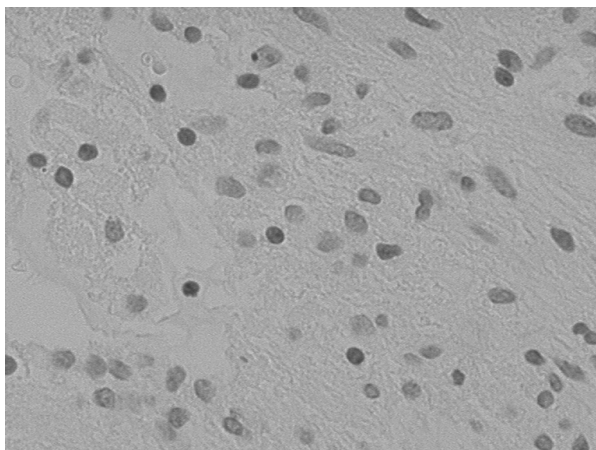


Fig. 4. H&E stained image ($768 \times 576 \times 8$ bit) of an astrocytic tumor ($\times 400$).

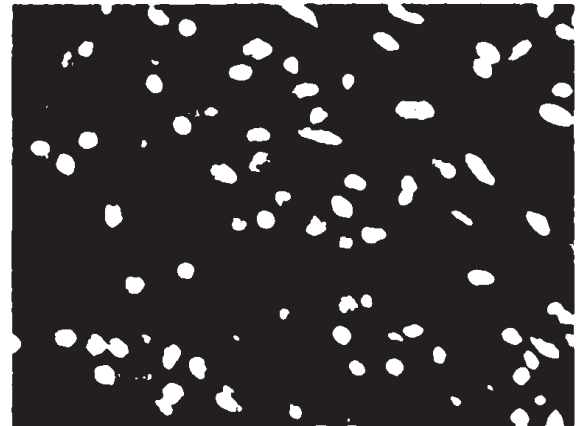


Fig. 5. Resulting binary image after applying the PNN clustering segmentation algorithm that discriminated pixels belonging to nuclei from pixels belonging to surrounding tissue.

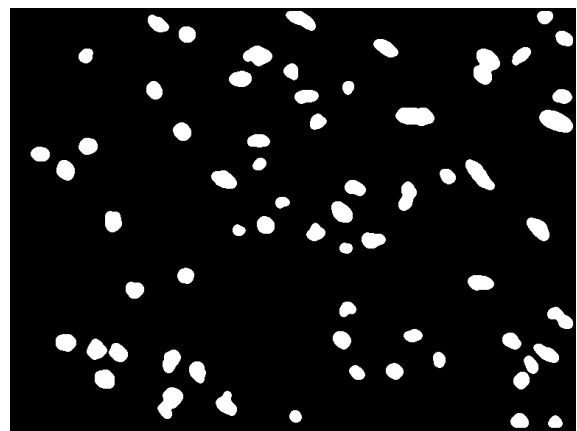


Fig. 6. The binary image was further processed with morphological filters, size filters (nuclei less than 200 pixels were omitted) and fill holes operations.

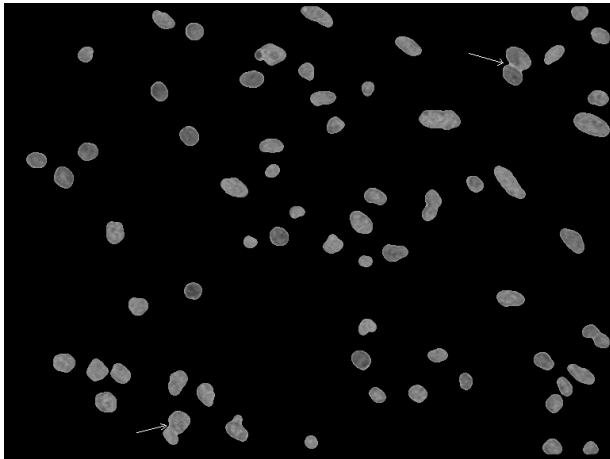


Fig. 7. Final segmentation was achieved by superimposing the original image to the morphological processed image. Arrows indicate closely spatially located nuclei that were incorrectly identified by the algorithm as one structure.

morphological-open filters, size filtering (nuclei less than 200 pixels were omitted) and fill-holes operations. Final segmentation was obtained by superimposing the original image to the morphological processed image (Fig. 7). According to the segmentation's algorithm evaluation procedure described in Sec. 2.2, the accuracy of correctly identified nuclei in randomly selected images ranged between 76.9–93.7%, with an average of 86.5% (see Table 2).

Concerning image classification, high-grade tumours were classified with an accuracy of 91.0%, low-grade tumours with 95.0% and 'suspicious' cases of intermediate grade with 88.3%. Overall accuracy was 92.1% (Table 3). From the 67 high-grade cases, 61 were correctly identified as high grade, 4 incorrectly as low grade and 2 as of intermediate grade. From the 61 low-grade cases, 58 were correctly classified as low grade, 1 incorrectly as high grade and 2 as of intermediate grade. Finally, from the 12 'suspicious' cases of intermediate grade, 10 were correctly identified as of intermediate grade and 2 incorrectly as high grade. Best feature vector combination for the discrimination of (i) low from high-grade tumours was energy, roundness, correlation and short run emphasis; (ii) low from intermediate grade tumours was concavity, angular second moment, roundness and energy, (iii) high from intermediate grade tumours was energy, area, gray level non-uniformity and concavity.

Table 2. Evaluation of the segmentation algorithm for 25-segmented H&E images of astrocytomas.

	Diagnosis	Accuracy
Image 1	Grade III	83.3
Image 2	Grade III	88.9
Image 3	Grade III	84.9
Image 4	Grade II	93.7
Image 5	Grade IV	88.6
Image 6	Grade IV	76.9
Image 7	Grade II	92.7
Image 8	Grade II	88.7
Image 9	Grade IV	79.1
Image 10	Grade III	81.5
Image 11	Grade IV	81.2
Image 12	Grade III	83.4
Image 13	Grade II	92.0
Image 14	Grade IV	81.3
Image 15	Grade III	84.0
Image 16	Grade III	84.6
Image 17	Grade IV	83.3
Image 18	Grade III	89.7
Image 19	Grade II	92.2
Image 20	Grade II	90.4
Image 21	Grade II	94.1
Image 22	Grade II	92.5
Image 23	Grade II	92.2
Image 24	Grade II	85.1
Image 25	Grade IV	79.7

4. Discussion

In this study a computer-assisted diagnosis system for astrocytomas grading was introduced by developing a novel PNN clustering algorithm for image segmentation and an SVM-based decision tree procedure for image classification.

Regarding image segmentation results, nuclei segmentation accuracy reached approximately 86.5%. Accurate segmentation of nuclei is of crucial importance to guarantee correct results in computer-assisted microscopy.²⁶ Nuclei encode significant diagnostic and prognostic information, that if quantified can potentially allow the prediction of the disease course. Previous studies that have investigated the demanding task of nuclei segmentation^{27–30} have reported relatively high segmentation accuracies, such as 85%,²⁷ 89%,²⁸ and 99%.²⁹ In Ref. 27 a method has been presented for the automatic segmentation of nuclei in breast fine needle aspiration images using the Hough transform. In Refs. 28 and 29 images from Papanicolaou stained

Table 3. Truth table demonstrating classification results for the 140 cases of astrocytomas.

Diagnosis	Low grade (grade I–II)	High grade (grade III–IV)	Suspicious cases (grade II–III)	Classification results (accuracy)
Low grade	58	1	2	95.0%
High grade	4	61	2	91.0%
Suspicious cases	0	2	10	83.3%
Overall accuracy				92.1%

smears have been automatically segmented using Neural Networks and active contours respectively. In Ref. 30 a method for segmentation of immunohistochemically stained nuclei has been presented using a supervised pixel based classification algorithm and watershed operations. However, these studies have greatly depended on *a priori* assumptions established for nuclei morphology and size. The proposed PNN clustering algorithm requires no assumptions concerning either nuclei size or shape. Additionally, it has to be stressed that in this study the routine H&E staining protocol was employed that is not as accurate in nuclei staining as other specialized staining protocols such as those used in previous studies (*ki-67*, Feulgen, Toluidine-blue and Papanicolaou staining).^{10–11,27–30}

The PNN-clustering algorithm is designed to identify different data partitions based on data distribution. In Ref. 31 a Gaussian mixture approach has been presented based on the data distribution; however this algorithm has required *a priori* assumptions on the data and noise distributions. In order to let the data “speak for themselves” the proposed PNN clustering algorithm was designed to make no *a priori* assumptions on data distribution, which was estimated by using a Gaussian Parzen kernel.³² Compared to Fuzzy C Means robust clustering methods,^{33–36} which have been designed to detect ellipsoid-shaped clusters and in some cases curves or surfaces, the proposed algorithm was more flexible and was able to detect clusters of arbitrary size and shape. Compared to NN-based robust clustering techniques,^{37–38} the PNN-based method was relatively faster since no feedback paths or iterative procedures were required. Finally, the algorithm needed some *a priori* information regarding the number of expected clusters, which in many cases does not present a problem.

Considering classification results, the method gave high rates in correctly discriminating low from

high grade tumours and ‘suspicious’ cases of intermediate grade with overall accuracy 92.1%. SVM with polynomial kernel of degree 2 optimized classification performance in both levels of the decision tree. The decision tree structure was preferred to using a single 3-class SVM classifier because it resembles the diagnostic procedure followed by the histopathologist and it is less complex and faster to implement. The number of support vectors indicated the good generalization capacity^{24–25} of the SVM classifier: (a) 17 for constructing the SVM classifier to discriminate low from high-grade tumours (13.2% of classified cases), (b) 10 to identify low from intermediate-grade tumours (14.7% of all classified cases), and (c) 12 to separate high from intermediate-grade tumours (15.6% of classified cases).

These results are promising compared to those presented in literature for automatic grading systems that have utilized the WHO scheme but more demanding staining procedures, such as 55% in Ref. 9, 66% in Ref. 12, 88% in Ref. 8 and 89.7% in Ref. 21. In contrast to previous studies,^{9–12,21} the proposed method was designed to additionally automatically specify “suspicious” cases, i.e., those cases that cannot be definitely characterized either as low or high grade (grade II to III cases). The difficulty for histopathologists to provide a definite diagnosis for cases between grade II and grade III might be explained due to that astrocytomas malignancy develops along a biological continuum.⁵ It is, thus, very important for an automatic grading algorithm to specify “suspicious” cases of intermediate grade that need re-examination because clinical management is strongly tied to grade assignment⁶: an incorrect assessment of a low grade tumour as high grade might lead to an excessively aggressive therapy, whereas a false assignment of a high grade tumour as low grade might lead to a less aggressive therapy. In our dataset 12 cases were characterized by the histopathologist as “suspicious” cases

(between low and high grade). The PNN algorithm correctly identified 10 of the 12 “suspicious” cases, with accuracy 83.3%. The remaining two cases were classified by the algorithm as high grade. The first of these 2 cases was a low-grade astrocytic tumour that recurred after 2 years from first diagnosis and surgical removal. The second case was an astrocytic tumour with histological characteristics of grade II; however, the histopathologist assessed this tumour as between grade II and grade III due to the existence of extended necrosis.

5. Conclusion

The method proposed in this study gave high accuracy in automatically separating low from high grade astrocytomas while specifying those ‘suspicious’ cases of intermediate grade that need re-examination. The method was developed to be compatible with the WHO grading system and H&E staining protocol, which are the most common standards in clinical routine.^{3,9} Therefore the proposed computer assisted microscopy system might be utilized in daily clinical practice without the need of performing specialized staining processes that increase the time and cost of diagnosis and without burdening histopathologists with the need to familiarize with new grading protocols.

Appendix A

In high-grade tumours the existence of giant nuclei with irregular curvature and shape is frequently pronounced. On the contrary, in low-grade tumours nuclei are smaller, round and look more or less alike. Thus, the quantification of morphological characteristics of nuclei may contain important diagnostic information concerning the grade of tumours. In this study 18 morphological features were computed. These features comprised measurements of area, roundness and concavity. Additionally, the mean value, maximum value, minimum value, standard deviation, skewness and kurtosis of each of the 3 morphological features were calculated.

Textural features have been shown in many studies as powerful descriptors of tumours malignancy.⁷⁻¹³ In this study three kinds of textural features were calculated: *First order statistical features* that comprised measurements of the mean value, standard deviation, skewness and kurtosis

from nuclei histogram. *Second order statistical features* from the co-occurrence matrix.¹⁹ Although the first order statistical features describe the gray level distribution of nuclei, these features do not give any information concerning the spatial distribution of the various gray levels inside nuclei. This type of information can be extracted from the co-occurrence matrix that describes the frequency of appearance of pairs of pixels at a distance d (inter-pixel distance) with gray values I_1 and I_2 inside each nucleus. The co-occurrence features that were calculated were the following:

1. Angular Second Moment $ASM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (p(i, j))^2$, where N_g is the number of gray levels in the image, $i, j = 1, \dots, N_g$, and $p(i, j)$ is the co-occurrence matrix. ASM describes image smoothness and takes minimum values for smooth textures nuclei.
2. Contrast $CON = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (p(i, j))^2 \right\}, |i - j| = n$. CON increases for high contrast nuclei. The factor n^2 is enhances big differences.
3. Inverse Different Moment $IDM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{p(i, j)}{1+(i-j)^2}$. IDM increases for low contrast nuclei due to the dependence on $(i - j)^2$.
4. Entropy $ENT = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p(i, j) \log(p(i, j))$. ENT is a measure of randomness and takes low values for smooth nuclei.
5. Correlation $COR = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (ij)p(i, j) - m_x m_y}{\sigma_x \sigma_y}$, where m_x, m_y, σ_x and σ_y the respective mean values and standard deviations of p_x and p_y . COR encodes the gray tones dependencies in nuclei.

These 5 co-occurrence matrix based textural features were calculated with inter-pixel distance $d = 1$ and $d = 3$.

- Second order statistical features from the run length matrix.²⁰ The run length matrix describes the frequency of appearance of a set of consecutive pixels (run) having the same gray value. The run length features that were calculated were the following:

1. Short run emphasis $SRE = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_r-1} \frac{r(i, j)}{j^2}}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_r-1} r(i, j)}$, where $r(i, j)$ is the run length matrix, N_g is the number of gray values, N_r is the largest possible run, $i = 1, \dots, N_g, j = 1, \dots, N_r$. SRE tends to

emphasize short runs due to the division with j^2 and takes high values for coarser nuclei.

2. Long run emphasis LRE = $\frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_r-1} j^2 r(i,j)}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_r-1} r(i,j)}$. LRE tends to emphasize long runs and is large for smoother nuclei.
3. Gray level non-uniformity GLNU = $\frac{\sum_{i=0}^{N_g-1} (\sum_{j=0}^{N_r-1} r(i,j))^2}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_r-1} r(i,j)}$. GLNU increases for nuclei having many runs of the same gray level value.
4. Run length non-uniformity RLNU = $\frac{\sum_{j=0}^{N_g-1} (\sum_{i=0}^{N_r-1} r(i,j))^2}{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_r-1} r(i,j)}$. RLNU takes low values for nuclei with homogeneous distribution of runs.
5. Run Percentage RP = $\frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_r-1} r(i,j)}{P}$ where P is the total possible number of runs in the nucleus image. This feature takes its lowest value for nuclei with linear structures.

Acknowledgments

We thank European Social Fund (ESF), Operational Program for Educational and Vocational Training II (EPEAEK II) and particularly the Program IRAK-LEITOS for funding the above work. J. Tohka's work was supported by the Academy of Finland under the grants nos. 204782 and 104834 and the NIH/NCRR grant P41 RR013642, additional support was provided by the NIH Roadmap Initiative for Bioinformatics and Computational Biology U54 RR021813 funded by the NCRR, NCBC, and NIGMS.

References

1. L. M. DeAngelis, Brain tumours, *New England Journal of Medicine* **344** (2001) 114–123.
2. W. Shapiro and J. Shapiro, Biology and treatment of malignant gliomas, *Oncology* **12** (1998) 233–240.
3. World Health Organization, International histological classification of tumours: Histological typing of tumours of the central nervous system, 2nd edn. (Berlin, Springer-Verlag, 1993).
4. K. Herfarth, S. Gutwein and J. Debus, Postoperative radiotherapy of astrocytomas, *Seminars in Surgical Oncology* **20** (2001) 13–23.
5. R. A. Prayson, D. P. Agamanolis, M. L. Cohen and M. L. Estes, Interobserver reproducibility among neuropathologists and surgical pathologists in fibrillary astrocytoma grading, *J. Neurological Sciences* **175** (2000) 33–39.
6. W. Coons, P. Jhonson, B. Sceithauer, A. Yates and D. Pearl, Improving diagnostic accuracy and interobserver concordance in the classification and grading of Primary Gliomas, *Cancer* **79** (1997) 1381–93.
7. P. Sallinen, Quantitative pathology in astrocytic tumours, PhD dissertation, University of Tampere (1999).
8. N. Reinhold and W. Schlote, Topometric analysis of diffuse astrocytomas, *Analytical and Quantitative Cytology and Histopathology* **25** (2003) 12–18.
9. C. Decaestecker, I. Salmon, O. Dewitte, I. Camby, P. Van Ham, J. Pasteels, J. Brotchi and R. Kiss, Nearest-neighbor classification for identification of aggressive versus nonaggressive astrocytic tumours by means of image cytometry-generated variables, *J. Neurosurgery* **86** (1997) 532–537.
10. M. Scarpelli, P. Bartels, R. Montironi, C. Galluzzi and D. Thompson, Morphometrically assisted grading of astrocytomas, *Analytical and Quantitative Cytology and Histology* **16** (1994) 351–356.
11. A. V. Wangenheim, G. H. Vince, H. Kolles, M. M. Richter and W. Feiden, Grading of Gliomas in stereotactic Biopsies with Neural Networks, in *Proc. 4th conf. Artificial Intelligence in Medicine*, Vol. 10 (1993), pp. 486–488.
12. N. Belacel and M. Boulassel, Multicriteria fuzzy assignment method: a useful tool to assist medical diagnosis, *Artificial Intelligence in Medicine* **21** (2001) 201–207.
13. P. Spyridonos, D. Cavouras, P. Ravazoula and G. Nikiforidis, Neural Network based segmentation and classification system for the automatic grading of histological sections of urinary bladder carcinoma, *Analytical and Quantitative Cytology and Histology* **26** (2002) 317–324.
14. O. Faugeras and W. Pratt, Decorrelation methods of texture feature extraction, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14** (1980) 323–332.
15. D. Glotsos, J. Tohka, J. Soukka and U. Ruotsalainen, A new approach to robust clustering by density estimation in an autocorrelation derived feature space, in *Proc. 6th NORDIC Signal Processing Symposium* (2004), pp. 296–299.
16. D. Specht, Probabilistic neural networks, *Neural Networks* **3** (1980) 109–118.
17. N. Piliouras, I. Kalatzis, P. Theocharakis, N. Dimitropoulos and D. Cavouras, Development of the Probabilistic Neural Network — Cubic Least Squares Mapping (PNN-LSM³) classifier to assess carotid plaque's risk, *Pattern Recognition Letters* **25** (2004) 249–258.
18. K. K. Yiu, M. W. Mak and S. Y. Kung, A comparative study on kernel-based probabilistic neural networks for speaker verification, *Int. J. Neural Systems* **12** (2002) 381–397.
19. R. M. Halarick, K. Shanmugam and I. Dinstein, Textural features for image classification, *IEEE Trans-*

- actions on Systems, Man and Cybernetics* **3** (1973) 610–621.
20. M. M. Galloway, Texture analysis using grey level run lengths, *Computer Graphics and Image Processing* **4** (1975) 172–179.
 21. D. Glotsos, P. Spyridonos, P. Petalas, D. Cavouras, I. Ravazoula, S. Dadioti, E. Lekka and G. Niki-foridis, Computer-based malignancy grading of astrocytomas employing a support vector machine classifier, the WHO grading system and the regular hematoxylin-eosin diagnostic staining procedure, *Analytical and Quantitative Cytology and Histology* **26** (2004) 77–83.
 22. V. Kechman, *Learning and Soft Computing* (MIT, USA, 2001), pp. 121–189.
 23. S. Theodoridis and K. Koutroubas, Pattern recognition, Academic Press **342** (1999).
 24. I. Kalatzis, D. Pappas, N. Piliouras and D. Cavouras, Support vector machines based analysis of brain SPECT images for determining cerebral abnormalities in asymptomatic diabetic patients, *Medical Informatics and the Internet in Medicine* **28** (2003) 221–230.
 25. A. Ifantis and S. Papadimitriou, The nonlinear predictability of the electrotelluric field variations data analyzed with support vector machines as an earthquake precursor, *Int. J. Neural Systems* **13** (2003) 315–332.
 26. T. Mouroutis, S. Roberts and A. Bharath, Robust cell nuclei segmentation using statistical modeling, *Bioimaging* **6** (1998) 79–91.
 27. K.-M. Lee and N. Street, A fast and robust approach for automated segmentation of breast cancer nuclei, in *Proc. Second IASTED Int. Conf. Computer Graphics and Imaging* (1999), pp. 42–47.
 28. S. J. McKenna, Automated analysis of papanicolaou smears, Ph.D. Dissertation, University of Dundee (1994).
 29. P. Bamford and B. Lovell, Unsupervised cell nucleus segmentation with active contours models, *Signal Processing* **71** (1998) 203–213.
 30. P. Ranefall, L. Egevad, B. Nordin and E. Bengtsson, A new method for segmentation of colour images applied to immunohistochemically stained cell nuclei, *Analytical Cellular Pathology* **15** (1997) 145–156.
 31. X. Zhuang, Y. Huang, K. Palaniappan and Y. Zhao, Gaussian mixture density modeling, decomposition, and applications, *IEEE Trans. Image Processing* **5** (1996) 1293–1302.
 32. E. Parzen, On the estimation of a probability density function and the mode, *Ann. Inst. Stat. Math.* **33** (1962) 1065–1076.
 33. H. Frigui and R. Krishnapuram, A robust competitive clustering algorithm with applications in computer vision, *IEEE Trans. Pattern Anal. Machine Intell.* **21** (1999) 450–465.
 34. P. W. Holland and R. E. Welsch, Robust regression using iteratively reweighted least squares, *Communication Statistics — Theory and Methods* **9** (1977) 813–827.
 35. W. Zheng and G. G. Yen, On computing the fuzzifier in JFlvq: A data driven approach, *Int. J. Neural Systems* **12** (2002) 149–157.
 36. H. Hamid Muhammed, Using weighted fixed neural networks for unsupervised fuzzy clustering, *Int. J. Neural Systems* **12** (2002) 425–434.
 37. P. Sykacek, Outliers and Bayesian inference, in *Proc. NC 98* (1998), pp. 973–978.
 38. S. Hawkins, H. He, G. Williams and R. Baxter, Outlier detection using replicator neural networks, *Neural Networks* (2002) 170–180.