CLASSIFICATION OF BRAIN TUMORS IN MR IMAGES

Qiang Wang[1,2] *Eirini Karamani Liacouras*[1,2] *Erickson Miranda*[1,2]
Uday S. Kanamalla[3] *Vasileios Megalooikonomou*[1,2]

[1] Department of Computer and Information Sciences, Temple University
[2] Data Engineering Laboratory (DenLab), Temple University
[3] Dept. of Diagnostic Imaging, School of Medicine, Temple University

ABSTRACT

We study the problem of classifying brain tumors as benign or malignant using information from magnetic resonance (MR) imaging and magnetic resonance spectroscopy (MRS) to assist in clinical diagnosis. The proposed approach consists of several steps including segmentation, feature extraction, feature selection, and classification model construction. Using an automated segmentation technique based on fuzzy connectedness we accurately outline the tumor mass boundaries in the MR images so that further analysis concentrates on these regions of interest (ROIs). We then apply a concentric circle technique on the ROIs to extract features that are utilized by the classification algorithms. To remove redundant features, we perform feature selection where only those features with discriminatory information (among classes) are used in the model building process. The involvement of MRS features further improves the classification accuracy of the model. Experimental results demonstrate the effectiveness of the proposed approach in classifying brain tumors in MR images.

1. INTRODUCTION

Early detection and classification of brain tumors is very important in clinical practice. Many researchers have proposed different techniques for the classification of brain tumors based on different sources of information [1, 2, 3]. In this paper we propose a process for brain tumor classification, focusing on the analysis of Magnetic Resonance (MR) images and Magnetic Resonance Spectroscopy (MRS) data collected for patients with benign and malignant tumors. Our aim is to achieve a high accuracy in discriminating the two types of tumors through a combination of several techniques for image segmentation, feature extraction and classification. The proposed technique has the potential of assisting clinical diagnosis.

Necessary preprocessing steps prior to characterization and analysis of regions of interest (ROIs) are segmentation and registration. Image registration is used to determine whether two subjects have ROIs in the same location. However, in this work we do not take into account the location of the tumor in the classification model so we do not employ registration. Image segmentation is required to delineate the boundaries of the ROIs ensuring, in our case, that tumors are outlined and labeled consistently across subjects. Segmentation can be performed manually, automatically, or semi-automatically. The manual method is time consuming and its accuracy highly depends on the domain knowledge of the operator. Extensive work on medical image segmentation (either automatic or semiautomatic) has been done; methods can be divided into two broad groups: those that incorporate prior spatial information and those that are solely signal-intensity based; reviews can be found in [4, 5, 6, 7]. Specifically, various approaches have been proposed to deal with the task of segmenting brain tumors in MR images [8, 9]. The performance of these approaches usually depends on the accuracy of the spatial probabilistic information collected by domain experts. In previous work [10], we proposed an automatic segmentation algorithm that is based on the fuzzy connectedness concept [11]. The main idea is to assign to every pair of voxels, x, y, in the image, a real number between 0 and 1 indicating their connectedness. Starting with several seed points, all the voxels are automatically assigned to the structure to which they have the highest connectedness value. Utilizing the statistical information cumulated during the segmentation process, this method can provide satisfying results even in cases where the boundaries of the ROIs cannot be easily identified.

Having segmented the ROI and in order to build a classification model, one needs to extract a set of discriminative features from the ROI. Most characterization techniques are based on extracted global visual features that refer to the entire image rather than to regions that are of interest. However, in medical images, feature extraction has to focus on specific regions and capture not only shape but also structural and internal volume properties that can be useful for building a classification model. Megalooikonomou et al. [12, 13] proposed a method that efficiently extracts a *k*-dimensional feature vector using concentric spheres in 3D (or circles in 2D) radiating out of the ROI's center of mass. The method has been applied successfully to classification and similarity searches of spatial ROIs.

In this paper, we propose an approach (see Figure 1) for building a classification model for brain tumors. Regions of interest (tumors) are first segmented from the MR images, and a group of features is extracted. Instead of employing all of the features to build the model, a preprocessing step of feature selection is performed aiming to remove the redundant features. Based on the statistical information, only the most informative features extracted from the MR images are utilized in the model building process. In addition, in this paper, we consider features from other sources (e.g., MRS data) in the classifier training process. This leads to improved classification accuracy.

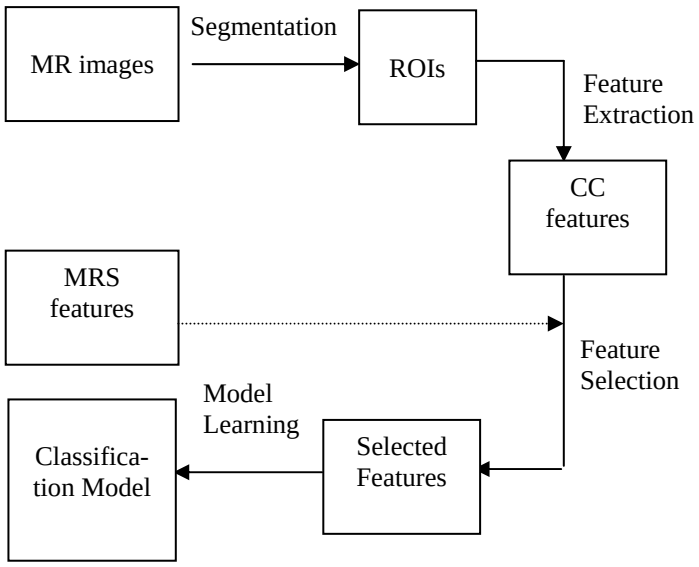


Figure 1. A framework for brain tumor classification

2. METHODOLOGY

There are four major steps in the proposed approach for brain tumor classification: (a) ROI segmentation: delineating the boundary of the tumor (ROI) in an MR image; (b) feature extraction: getting meaningful features of the ROI identified in the previous step; (c) feature selection: removing the redundant features; (d) classification: learning a classification model using the features.

2.1 Segmentation

We apply a segmentation algorithm (Fuzzy-Connectedness segmentation (FCS)) based on the concept of fuzzy connectedness to detect all the pixels on the MR images belonging to the tumor and discriminate from normal tissue. The main idea of fuzzy-connectedness is to assign to every pair of voxels, *x*, *y*, in the image, a real number between 0 and 1. This number is the fuzzy-connectedness of *x* to *y* and is utilized to denote the strength of the link between *x* and *y*. The segmentation algorithm begins with a given set of seed points, having at least one seed point for each type of

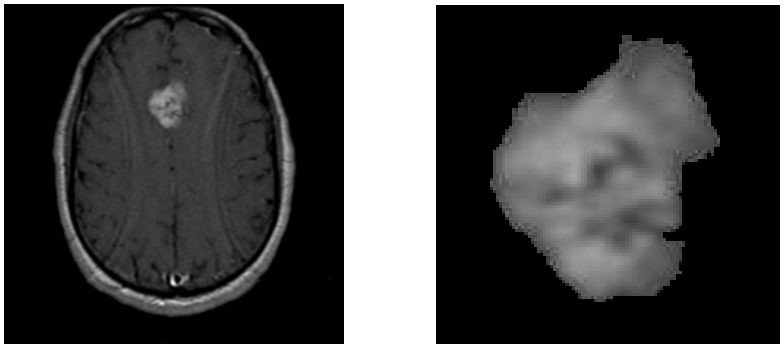
tissue. The selection of seed-points is more effective using some prior knowledge about the domain; they can be either manually chosen by an expert or automatically selected by a computer algorithm. The automatic segmentation process utilizes the strength of fuzzy-connections between points to construct a structure. Each point is assigned to the structure having a neighboring point with the highest fuzzy-connectedness value. The strongest connection is detected first, and the process repeats until the weakest connection is calculated. In that sense, the grade of membership in an object of an arbitrary point is its fuzzy-connectedness to a pre-defined seed point. A sequence of points is called a chain, where its links are the ordered pairs of consecutive points in the sequence. The strength of a chain is the length of its weakest link. The segmented object is defined by the number of points that are connected through a chain to the selected seed point of the object. Here, we give several definitions related to Fuzzy-Connectedness segmentation (FCS). More details can be found in [10].

Definition 1: For a positive integer K, a *K -FCS* of an image D is a function which maps each voxel $c \in D$ into a (K+1)-dimensional vector $\sigma^c = (\sigma_0^c, \sigma_1^c, \sigma_2^c, ..., \sigma_K^c)$, such that $\sigma^c \in [0,1]$, and for at least one k, $1 \le k \le K$, $\sigma_0^c = \sigma_k^c$; for $1 \le m \le K$ and $m \ne k$, $\sigma_m^c = 0$ or $\sigma_m^c = \sigma_k^c$.

Definition 2: A fuzzy affinity on D is a function : $\varphi : D^2 \to [0,1]$. φ (c, d) is the strength of the link (c,d) .

Definition 3: A chain in $U \subseteq D$ from c_0 to c_N is a sequence $C^{(0,N)} = (c_0, c_1, c_2, \ldots c_N)$ in U and the strength of the chain $\varphi_U(C^{(0,N)}) = \min(\varphi_U(c_{n-1}, c_n))$, $1 \le n \le N$.

An example of the segmentation of a tumor in Figure 2(a) is shown in Figure 2(b).



(a) (b)

Figure 2. (a) An original 2D axial MRI slice and (b) the segmented tumor (ROI) obtained using fuzzy-connectedness segmentation

2.2 Feature Extraction

Before the classification model can be built, meaningful features of the ROIs delineated during the process of segmentation, need to be extracted and used as input in the model learning process. The feature extraction we used in our framework was introduced in [12] and has been successfully applied to medical images [14]. The basic idea of this technique is to construct a series of 1,…,k concentric circles with regular increments of their radius radiating from the ROI's center of mass. For each radius increment, the fraction of the region occupied by the circle or the fraction of the circle occupied by the region is measured. In both cases, a feature vector consisting of k attributes is formed to represent the ROI. Based on the assumption that different classes of brain tumors tend to have different internal density characteristics and growth models, this intuitive method provides discriminative features. Figure 3 shows an example of the extraction of features using the concentric circle approach from a tumor.

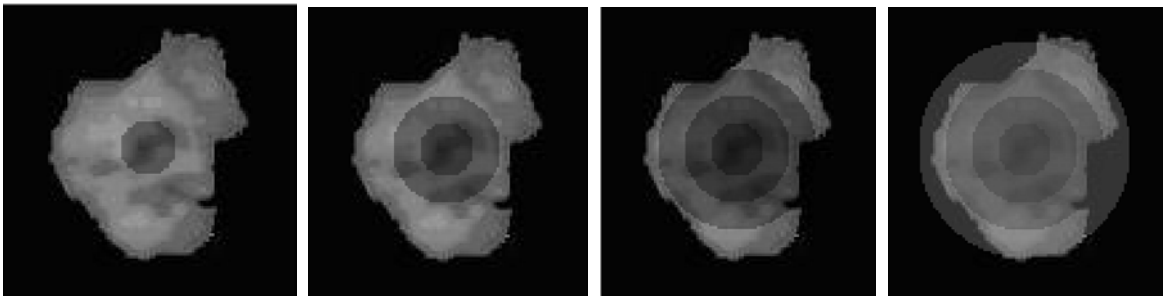


Figure 3. A snapshot of the characterization process of a tumor (ROI) using intersections with concentric circles

While the features extracted using the concentric-circle approach from MR images provide very useful information for classification, their effectiveness may be affected by some factors such as image quality and segmentation error. The involvement of other patient data may further improve the system's performance in classification.

Magnetic resonance spectroscopy (MRS) measures the concentrations of different chemical components within tissues. As chemical components such as n-acetylaspartate (NAA), Choline (Cho) and Creatinine (Cr) are thought to be markers of neuronal integrity, patterns of metabolic peak height ratios observed by MRS can further assist in clinic interpretation of MR images. In the proposed approach we fuse features extracted from MRI and MRS to better discriminate between benign and malignant tumors.

2.3 Feature Selection

To remove redundant features, a feature selection pre-processing step is performed before the actual model learning. As a commonly used process in the practice of machine learning, feature selection becomes necessary when we have only limited data samples while the number of features is very large. This is due to the fact that the number of samples required to learn a mulitiple dimensionality distribution increases exponentially as the number of features increases linearly. There are many approaches available for feature selection. In our case, we use the simple but very effective *Information Gain* criterion to evaluate the usefulness of each attribute in a feature vector. As a concept coming from information and decision tree theory, Information Gain shows the power an attribute has to discriminate among different classes.

Given class label Y and an attribute X, the information gain of Y given X is defined as:

$$\text{InfoGain}(Y|X) = H(Y) - H(Y|X)$$

where H(Y) and H(Y|X) are Y's entropy and conditional entropy, respectively. Interested readers are encouraged to check [15] for more details.

To make the selection process more robust, the data is first divided into several groups and the importance of the attributes is evaluated in each group. Finally the most discriminating features are selected combining the evaluation results from all groups. Note that some classification models may have their own inherent ability for feature selection, but this fact should not eliminate the usefulness of a feature selection preprocessing step. In general, concentrating on more informative features increases the robustness of the trained model.

2.4 Model Learning

There are many techniques available for constructing a classification model. Neural networks, decision trees and Bayesian networsk are among the most popular choices. While each technique has its strengths and constraints, in this work, we choose to utilize the decision tree model because of its nice interpretability.

A decision tree is a tree structure where leaf nodes represent classifications while branches show the related features leading to the classifications. Besides the ease of understanding, the decision tree model provides several other advantages over the other machine learning methods: it can deal with both nominal and categorical data; it is possible to

be validated with statistical tests; its training time is short and it is scalable to large datasets. All these characteristics are ideal for the analysis of medical data. When building a decision tree for a given dataset, all the features involved in the learning process are selected in the order of their discriminative power to add new branches to the tree until all the data samples under a node have the same class label. While a complex tree may fit perfectly the data samples used for training, it tends to provide a poor performance on new instances. This effect is called overfitting. In order to avoid overfitting, after a tree is built, usually a post-processing step of pruning is necessary to make the tree simpler and make the trained model more reliable and robust.

3. EXPERIMENTAL RESULTS

In our experiments, we used a dataset consisting of 50 MR brain images: 25 with benign tumors and 25 with malignant tumors. MRI was performed at 1.5 tesla and images consisted of unenhanced T1-weighted spin-echo [repition time (TR)/time to echo (TE)/number of excitations (NEX): 400/16/1], followed by fast spin-echo (FSE) T2-weighted (2800/90/1) sequences. The matrix was 256x128 for T1-weighted images and 256x160 for T2-weighted images. A 22cm field of view and a slice thickness of 5 mm with a 1.5 mm interslice gap were used with all imaging sequences. The axial T1-weighted sequence was then repeated following intravenous administration of 0.1 mmol/kilogram of intravenous gadolinium. The gadolinium-enhanced T1-weighted sequence was used for localization of MR spectroscopy voxel. The MRS was performed utilizing the multi-TE technique, consisting of short (TR/TE/excitations: 1150/35/128), intermediate (1250/144/128), and long (1350/288/128) TE. The size of the MRS localizer voxel was selected at our discretion, depending on the size of the tumor; its volume ranged from 4.0 – 8.0 cm^3. The voxels were typically placed over the most homogeneous solid enhancing portion of the tumor. Automated spectral processing was performed using the commercially available GE SAGE software and consisted of zero filling and Fourier transformation of the free induction decay signal followed by zero and first order phasing and baseline correction of the frequency spectra. The peak area concentrations of various metabolites were determined. The areas under the curves were calculated for myo-inositol (mI) at 3.5 ppm, choline-containing compounds (Cho) at 3.2 ppm, creatine (Cr) at 3.0 ppm, glutamate/glutamine (Glx) at 2.2-2.4 ppm, n-acetylaspartate (NAA) at 2.0 ppm and lipids at 0.9 (Lip1) and 1.3 (Lip2) ppm.

We first used the FCS method to detect the regions of interest (tumors) in MR images. Due to the variability of the size of ROIs in different images, we normalized the result of segmentation by using a background template with the same dimensionality: 100 x 100 pixels. The normalization of image size is important since it guarantees that the content-based features (concentric circle features in our case) extracted out of different images are comparable to each other. During the feature extraction step, we used as increment of radius 6 pixels to obtain a concentric-circle feature vector of 16 attributes for each tumor (since each subject in our study had one tumor this feature vector corresponds to a subject).

The 16 attributes in the feature vector may not have the same amount of discriminative information due to several reasons: (a) the choice of the increment of radius may not be a perfect fit for the purpose of feature extraction, and may lead to the presence of attributes that are not very useful, (b) the background pixels introduced by normalization make the last several attributes have less discriminative information. Using the feature selection pre-processing step, 4 out of 16 features [CC1, CC3, CC6, CC7] were shown to have most discriminatory information and they were used to build the decision tree.

To build the classifier, we used the C5.0 algorithm [16]. This is a commercial version of C4.5 decision tree algorithm and is now widely utilized in many classification tasks. Compared to previous decision tree algorithms, C5.0 provides better classification accuracy and more efficient computation.

To make our results more reliable, a 5-fold cross validation was performed in our experiments: each time 80 percent of the data were used for training the classifier while the other 20 percent acted as testing data. In the first experiment, we used as input only the four selected concentric-circle features [CC1, CC3, CC6, CC7] of the MR images. As shown in Table 1, the average classification accuracy achieved was 69%.

As discussed in Section 2.2, several factors may affect the effectiveness of the concentric-circle features. In order to further improve the classification accuracy, we added four more features obtained from MRS of the same patients showing the concentrations of chemical components within tissues: Cr-area, bGlx-area, lip1.3-area and a-glx-area. In order to fuse these two groups of attributes and avoid the dominance of any specific feature, we normalized the data to

make all the attributes have the same range: [0 1]. As shown in the table, the average classification accuracy increased to 79%. Figure 4 shows the actual decision tree generated with the C5.0 algorithm. Note that not all the attributes are

Table 1. Classification results

	Avg Accuracy	Std. Deviation
MRI features	69.0%	6.0%
MRI and MRS features	79.0%	6.4%

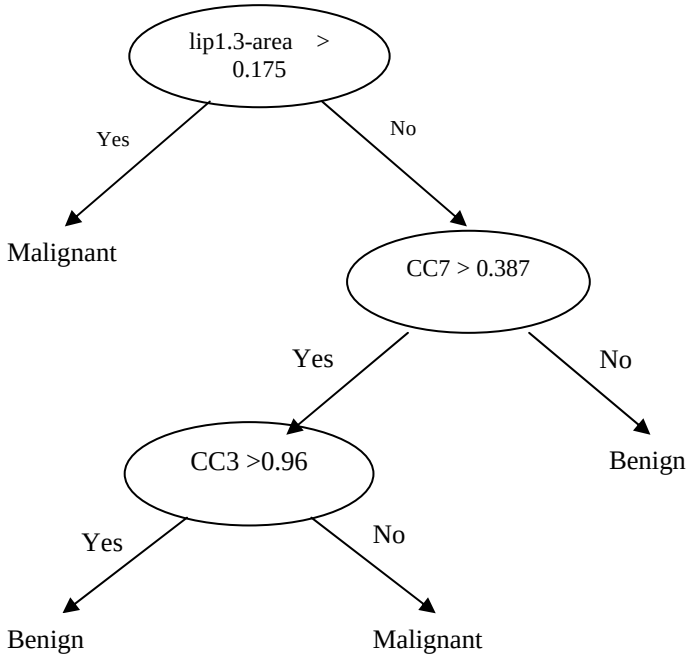


Figure 4. Decision tree with MRI and MRS attributes

involved in the tree. With C5.0's pruning ability, the decision tree tends be as simple as possible.

4. DISCUSSION AND CONCLUSIONS

In this paper, we propose an approach for classifying brain tumors in MR images. The purpose is to develop tools for discriminating malignant tumors from benign ones assisting decision making in clinical diagnosis. The proposed approach utilizes a combination of different techniques and is composed of several steps including segmentation, feature extraction and model learning. We also demonstrate that the fusion of data from other sources (MRS data in our case) can further improve the system's performance and provide encouraging results. Even though in this paper, we discuss mainly the application of the proposed approach to 2D MR images, it can be easily extended to 3D MR volumes, since both the segmentation and feature extraction techniques have been successfully applied to 3D volumes. Future work includes incorporating more specific information, such as the spatial information about the tumor into the

classification model. As MRS attributes show different values at different locations, the spatial information is expected to help in further improving the classification accuracy.

6. ACKNOWLEDGEMENTS

This project was supported by NIH research grant R01 MH68066-04 funded by the National Institute of Mental Health and the National Institute of Neurological Disorders and Stroke and the National Institute on Aging and by NSF research grant IIS-0237921.

7. REFERENCES

[1] R. Silipo; G. Deco, and H. Bartsch. “Brain tumor classification based on EEG hidden dynamics”. Intelligent Data Analysis, Volume 3, Number 4, pp. 287-306(20), October 1999.

[2] C. Majós, M. Julià-Sapé, J. Alonso, M. Serrallonga, C. Aguilera, J. Acebes, C. Arús and J. Gili. “Brain Tumor Classification by Proton MR Spectroscopy: Comparison of Diagnostic Accuracy at Short and Long TE”. American Journal of Neuroradiology 25:1696-1704, November-December 2004.

[3] L. Lukas, A. Devos, J.A.K. Suykens, L. Vanhamme, F.A. Howe,C. Majo´s, A. Moreno-Torres, M. Van Der Graaf, A.R. Tate, C. Aru´s, and S. Van Huffel. “Brain tumor classification based on long echo proton MRS signals”. Artificial Intelligence in Medicine: 31, 73—89, 2004.

[4] N. Pal and S. Pal, "A review on image segmentation techniques," Pattern Recognition, vol. 26, pp. 1277-1294, 1993.

[5] Y. Zhang, "A survey on evaluation methods for image segmentation," Pattern Recognition, vol. 29, pp. 1335-1346, 1996.

[6] A. Worth, N. Makris, V. Caviness, and D. Kennedy, "Neuroanatomical segmentation in MRI: Technological objectives," International Journal of Pattern Recognition and Artificial Intelligence, vol. 11, pp. 1161-1187, 1997.

[7] T. McInerney and D. Terzopoulos, “Deformable Models in Medical Image Analysis: A Survey”, *Medical Image Analysis*, 1(2):91–108, 1996.

[8] M. Prastawa, E. Bullitt, N. Moon, L. Van Leemput, and G. Gerig. “Automatic brain tumor segmentation by subject specific modification of atlas priors”. Acad. Radiology. 10:1341-1348, 2003.

[9] M. Schmidt, I. Levner, R. Greiner, A. Murtha, and A. Bistritz. "Segmenting Brain Tumors using Alignment-Based Features". International Conference on Machine Learning and Applications, Los Angeles, December 2005.

[10] D. Kontos, Q. Wang, V. Megalooikonomou, A. H. Maurer, L. C. Knight, S. Kantor, R. S. Fisher, H. P. Simonian, H. P. Parkman, "A 3D Image Analysis Tool for SPECT Imaging", Proceedings of the SPIE Conference on Medical Imaging: Visualization, Image-Guided Procedures and Display, Vol. 5744, pp. 839-847, San Diego, CA, Feb. 2005.

[11] A. Rosenfeld, “Fuzzy Digital Topology”, *Information and Control*, vol. 40, pp. 76-87, 1979.

[12] V. Megalooikonomou, H. Dutta, D. Kontos, "Fast and Effective Characterization of 3D Region Data", Proceedings of the IEEE International Conference on Image Processing (ICIP), Rochester, NY, pp. 421-424, Sep. 2002.

[13] D. Kontos and V. Megalooikonomou, "Fast and effective characterization for classification and similarity searches of 2D and 3D spatial region data", *Pattern Recognition*, Vol. 38, No. 11, pp. 1831-1846, 2005.

[14] D. Kontos and V. Megalooikonomou, "Fast and Effective Characterization of 3D Region of Interest in Medical Image Data", Proceedings of the SPIE International Symposium on Medical Imaging 2004, San Diego, CA, Feb. 2004, Volume 5370 Medical Imaging 2004, pp. 1324-1331.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, 1991.

[16] Rulequest Research, “C5.0: An Informal Tutorial”, http://www.rulequest.com/see5-unix.html.