

PROBABILISTIC SEGMENTATION OF BRAIN TUMORS BASED ON MULTI-MODALITY MAGNETIC RESONANCE IMAGES

Hongmin Cai, Ragini Verma, Yangming Ou, Seung-koo Lee, Elias R. Melhem, Christos Davatzikos

Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania

ABSTRACT

In this paper, multi-modal Magnetic Resonance (MR) images are integrated into a tissue profile that aims at differentiating tumor components, edema and normal tissue. This is achieved by a tissue classification technique that learns the appearance models of different tissue types based on training samples identified by an expert and assigns tissue labels to each voxel. These tissue classifiers produce probabilistic tissue maps reflecting imaging characteristics of tumors and surrounding tissues that may be employed to aid in diagnosis, tumor boundary delineation, surgery and treatment planning. The main contributions of this work are: 1) conventional structural MR modalities are combined with diffusion tensor imaging data to create an integrated multimodality profile for brain tumors, and 2) in addition to the tumor components of enhancing and non-enhancing tumor types, edema is also characterized as a separate class in our framework. Classification performance is tested on 22 diverse tumor cases using cross-validation.

Index Terms—Brain tumors, tissue classification, multi-modal MRI data, edema.

1. INTRODUCTION

Effective brain tumor treatment ideally calls for an accurate identification of boundaries between tumor, edema and healthy tissue. This is very challenging mainly owing to the fact that high-grade tumors are inherently diffuse and infiltrative, they invade the surrounding healthy tissue, and are heterogeneous, comprising enhancing and non-enhancing tumor tissue types and edema, rendering the transition from tumor to healthy tissue gradual. It is therefore challenging, if possible at all, to identify a clear transition from healthy tissue to edema to tumor by an inspection of the MR images alone. Clinically, conservative treatments based primarily on clearly visible tumor leave large parts of brain tissue untreated, likely leading to faster tumor recurrence and spread, and lower chance of survival. This paper aims at creating tissue profiles that identify different tumor components, edema and healthy tissue using a combination of several structural MR modalities and diffusion tensor MRI.

While clinical decisions on tumor treatments rely, in part, on radiological evaluation of structural images, such as Fluid Attenuated Inversion Recovery (FLAIR) and T1-weighted MR images, to obtain estimates of tumor, edema and healthy tissue, that may be rater dependent, several automated methods of tumor segmentation [1-4] have produced promising results mostly in differentiating tumor and normal tissue based on the traditional T1 and/or T2 MR modalities. However, perhaps due to the lack of variability in the information captured in the MR modalities that provide distinctive appearance signature of each tissue type, most of the existing methods have difficulty in differentiating tumor components and edema. In addition, although diffusion tensor imaging (DTI) has been successfully employed to investigate the tumor progress along the white matter (WM) tracts [5], it has never been integrated with structural modalities.

In this paper, we seek to address and alleviate these issues by combining structural MRI and DTI images into a multimodality tissue profile, which paves the way for classifying healthy and tumor tissues, followed by a categorization of brain tissue into more specific classes of enhancing tumor (ET), non-enhancing tumor (NET), edema, white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF). The proposed brain tissue classification framework incorporates intensities from each modality into an appearance signature of each voxel and trains appearance based classifiers using a combination of pattern classification techniques, based on training samples identified by a human expert. In addition to the hard segmentation of tumor components, a probability map is also generated for each of the six classes that characterize potential tumor infiltration and tumor growth prediction. The contributions of this work are: 1) creation of a multimodality tumor profile by integrating DTI images with conventional structural images, using tumor data from several patients; 2) investigation of the potential of this multi-modal classification in differentiating edema from the tumor components. Accurate and consistent tumor classification results for several tumor brains illustrate the robustness of our framework, and suggest potential applications in assessing tumor growth and in computer-guided surgery.

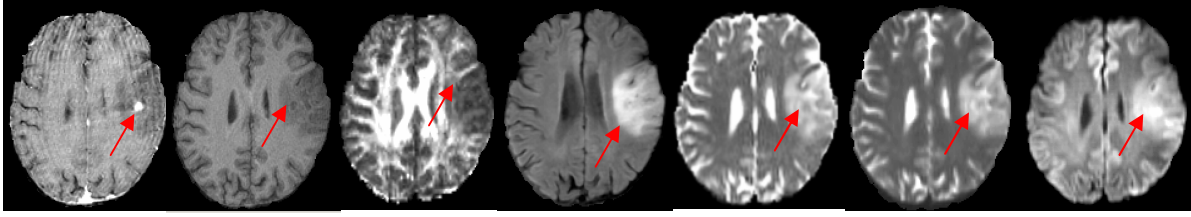


Fig. 1. A representative slice from each of the seven MR modalities used in creating the multimodality tissue profile. These have been rigidly co-registered to the T1 image of the patient. From left to right, the images are GAD, T1, FA, FLAIR, ADC, B0 and DWI. We use red arrows to stress tissue differences across the MR modalities.

2. METHOD

With the aim of distinguishing between healthy tissue and tumor components, our classification strategy defines 6 types of tissue classes: tumor (ET and NET), healthy tissues (WM, GM and CSF), and edema. Based on expert-defined training samples, classifiers were trained for each of the tissue types using information from a single patient or by pooling training data from several patients leading to intra- and inter-patient framework and were applied to new data from the same or another patient. Details of our framework are provided below.

2.1 Data Acquisition

For creating our multi-modality profile, we use seven MR images: five structural MR acquisition protocols, namely, B0, Diffusion Weighted Images (DWI), Fluid-Attenuated Inversion Recovery (FLAIR), T1-weighted, and gadolinium-enhanced T1-weighted (GAD), and two scalar maps computed from the DTI, namely, Fractional Anisotropy (FA) and Apparent Diffusion Coefficient (ADC).

2.2 Preprocessing

Prior to creating the intensity features from these images, the images are skull stripped and Gaussian smoothed using FSL [6]. Then, for each patient, all the modalities are rigidly co-registered to the T1-weighted image using FSL's registration algorithm, called FLIRT[7]. It may be noted that as the feature vectors are created by fusing information across modalities from within the same patient, rigid registration suffices between the modalities. In order to combine training samples from different patients, the images of the same modality are histogram matched across all patients. Fig. 1 shows representative slices from each of the acquisition protocols used to define the intensity feature vector for each voxel.

2.3 Definition of Training Samples

In order to train a robust classifier for each tissue class, we require samples of ET, NET and edema based on expert knowledge. Training samples for each of these classes were conservatively identified by a neuro-radiologist (SKL) typically using the FLAIR and GAD-T1 images. Edema is very difficult, if possible at all, to define with high confidence, because it is often mixed with infiltrating tumor.

In defining edema, our neuro-radiologist selected regions that based on the inspection of several MR modalities like GAD (for enhancing tumor) and FLAIR (for determining tumor boundaries). This was combined with implicit spatial knowledge about proximity of abnormal tissue to tumor, which would be identified as edema. Training samples for the healthy tissue (WM, GM and CSF) classes were defined using segmentation. After masking out a large area that contains the tumor and possibly, some normal tissue, a k-means clustering based method, is applied to segment the remaining brain regions that are purely normal into WM, GM and CSF sub-regions. Furthermore, to avoid bias in the training phase, the number of voxels selected in each WM/GM/CSF sub-region is set equal to the average number of samples in enhancing, non-enhancing and edema classes. See Fig. 2 (a – c) for training samples identified in red.

2.4 Design of Features

The *feature vector* for each voxel $\bar{x} \in I$, where I is 3D image volume, is defined as

$$\bar{v}_{\bar{x}} = [I_{\bar{x}}^{(ADC)}, I_{\bar{x}}^{(B0)}, I_{\bar{x}}^{(DWI)}, I_{\bar{x}}^{(FA)}, I_{\bar{x}}^{(FLAIR)}, I_{\bar{x}}^{(T1)}, I_{\bar{x}}^{(GAD)}]^T$$

where $I_{\bar{x}}^{(M)}$ denotes the intensity of image of modality M at voxel \bar{x} . These feature vectors are defined at each voxel in the training samples. In order to incorporate the variability around a voxel, we extend this voxel-wise feature to incorporate neighborhood information by using four of its neighbors. 7 dimensional intensity features for these 5 voxels are stacked into a long vector (35 dimensional), which is then used as a feature vector.

2.5 Classifier Construction

We construct two kinds of classifiers: 1) Intra-patient classifier: classifier is built using only half of each patient's expert defined training sample, then tested on the remaining half and 2) Inter-patient classifier: the classifier is trained and tested on separate datasets. Because our database is quite limited at this point, we used leave-one-out cross-validation mechanism. At the outset, it may be said that intra-patient classification is good in cases for which conservative training samples can be identified on the patient. Inter-patient classification addresses new cases for which no training data is available.

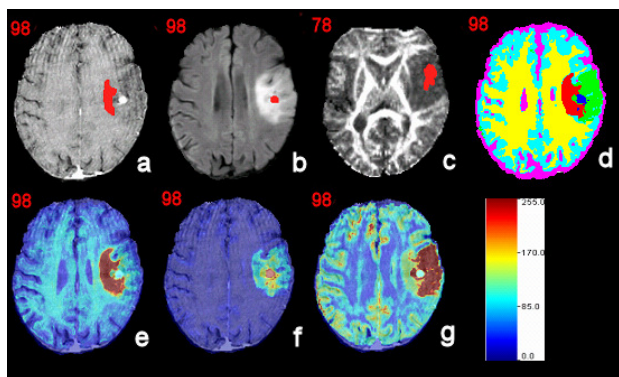


Fig. 2. Intra-patient segmentation on slice 82 of a typical patient with. (a) the edema training samples (red) overlaid on GAD; (b) the enhancing tumor training samples overlaid on FLAIR image; (c) the non-enhancing tumor training samples overlaid on FA image; (d) the hard segmentation results, in which class memberships are binarized through a threshold (e.g. 50%); (e-g) the probability maps for edema, enhancing and non-enhancing tumor.

Intra-patient classification: We use the Quadratic Discriminant Analysis (QDA) [8] method, to design discriminant functions for each of the 6 tissue classes. By computing the mean and the covariance matrix over the feature vectors of the training samples for the 6 tissue classes, we obtain a Quadratic Discriminant function [8] for each tissue class, that we refer to as the respective *tissue class classifier*. This discriminant evaluated at each voxel, provides the posterior probability of that voxel belonging to one of the 6 classes of ET, NET, edema, WM, GM and CSF. This produces a voxel-wise probability map for each brain, one pertaining to each of the 6 tissue classes. These discriminant values are normalized for visualization purposes. In addition, we can obtain hard segmentation by assigning the voxel to the class having the highest discriminant value, among the six classes. By assuming multivariate Gaussian distribution, the discriminant function can be computed efficiently and provides fast and efficient classification. The classifiers were trained on half of the training regions for that patient and tested on the remaining half.

Inter-patient classification: Classification of tumor and healthy tissue of a patient has a high accuracy in our framework when tested on that same subject. While useful for individual patient analysis and for treatment planning, such a profile can only be applied to current and perhaps future scans of that patient only, owing to the fact that the profile will not be able to capture the variability across patients. Indeed, these intra-patients classifiers typically fail on new patients, owing to the tumor variability. This motivated the definition of classifiers for tissue types, using training data from many different patients, incorporated into

an Support Vector Machine (SVM)-based framework. In this case, we combine training samples from across subjects, to obtain a more generalizable tissue classification. We design an SVM based classifier for all the 6 tissue classes by taking training samples from all the patients [9]. Due to the high variability across individuals, Quadratic Discriminant classification with their multinomial assumption does not provide adequate classification. We define classifiers, one pertaining to each of healthy (WM, GM and CSF combined), ET, NET and edema, in a one-versus-all framework. As SVM classifiers are tolerant to high variability, a single class for healthy tissue suffices and in addition, data from several patients can be combined to obtain robust classifiers. Responses from the classifiers are combined into a voting framework to obtain tissue classification. The classifiers are validated using a leave-one-out mechanism on the patients, that is, classifiers were trained using training samples from all patients except one, which was used for testing. We now proceed to apply our framework to a dataset of tumor patients.

3. RESULTS AND DISCUSSION

The experiments were conducted with the aim of distinguishing between tumor tissue types in patients and identifying regions of probable abnormality by using the multi-modality tumor profile. We have applied our 1) intra-patient and 2) inter-patient classification on a dataset consisting of twenty-two patients with newly diagnosed primary brain tumors who have not received any therapy prior to imaging at our institution (Hospital, University of Pennsylvania). All these patients have been diagnosed with a high grade (grade 3 or 4) tumor. The MR data for each patient was acquired either on a 3T Scanner (Siemens, Trio) or on a 1.5T (GE Medical Systems, Genesis Trio) scanner, under an IRB approved protocol with informed consent and is HIPAA compliant.

3.1. Intra-patient classification

For each patient in our dataset, we use half of the ground truth for training Quadratic Discriminant classifiers and the remaining (spatially non-contiguous) half region for testing each of the 6 tissue classifiers. The multivariate discriminant is computed for each voxel in the brain using Quadratic Discriminant classifiers, as explained in Section 2.5, and then normalized into a brain tissue probability map.

Fig. 2 shows the results of an application of our framework on one typical patient – female, age 53, diagnosed with a grade 3 tumor: Oligodendroglioma. Figure 2 shows representative training samples (red) for (a) edema overlaid on a GAD slice; (b) enhancing tumor training sample overlaid on FLAIR image; and (c) non-enhancing tumor overlaid on an FA image slice. Half of such training samples are used to create classifiers for these three classes. Fig. 2(d) gives the hard segmentation results, in which class

memberships are strictly binarized through a threshold. The classification accuracy rates were 98.75% for edema, 96.53% for ET, and 96.73% for NET, with the healthy tissue (GM, WM and CSF) classified with 99.85% accuracy on expert defined training regions. After grouping ET, NET and ED into an abnormal class, and GM, WM and CSF into a normal class, a sensitivity of 97.65% and specificity of 99.78% is obtained for the patient. Fig. 2(e-g) show the probability maps for edema, ET, and NET, respectively. The probability maps of the tissue types, generated by their respective classifier helps characterize the heterogeneity of the tissue, since each voxel gets a vote from all the classes. Thus each voxel not only gets characterized as healthy or

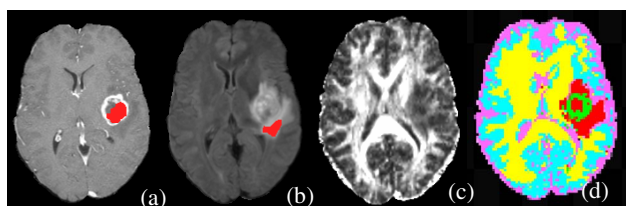


Fig.3. Across-patient segmentation on slice 85 of patient, From left to right, the images are: (a) the enhancing tumor training samples (red) overlaid on GAD; (b) the edema training samples overlaid on FLAIR image; (c) FA image; (d) the hard segmentation results, in which class memberships are obtained from SVM voting.

unhealthy but also if it is unhealthy, the extent to which it is a voxel from an enhancing or non-enhancing tumor type or edema. This is a significant contribution since tumors are highly heterogeneous and knowing their composition will help target the treatment of these regions better. We obtain similar high classification rates for each of the 22 patients.

3.3. Across-patient classification

We create 6 tissue classifiers using expert defined training samples from several patients, as explained in section 2.5. Figure 3 shows the classification by combining responses from 6 one-versus-all SVM classifiers applied to a tumor patient. We obtain classification rates of 98% for the edema and 73% for the enhancing-tumor on the training samples identified by the expert. Retrospectively examined, some likely errors in the gold standard definition seemed to explain the relatively low classification rate for enhancing tumor. It may be noted that we use training samples from all but this patient to produce the classifiers, and test the classifier on this left-out patient. In comparison, when Quadratic Discriminant classifiers were used to train across-patient classifiers, this patient gave very low classification rate, likely because probability density functions are not unimodal Gaussian.

4. CONCLUSIONS

This study applied a multivariate nonlinear classification scheme to the problem of soft tissue segmentation in brain

tumor patients. It combines conventional structural MRI with DTI, and used them to train classifiers for the tumor types of enhancing and non-enhancing tumor, edema and healthy tissue. The accurate distinction of the tumor tissue from healthy tissue as shown in Figs. 2(d) and Fig. 3 indicates that the framework can be useful in integrating multi-modality information into a combined profile and use it for classification. The hard segmentation as well as the probability maps can potentially provide a better understanding of the spatial distribution of healthy tissue, tumor and edema, thereby assisting in treatment or surgical planning. In the future, we plan to incorporate texture information into our features to distinguish between high grade and low grade tumors. In addition, we propose to build a two stage framework, in which SVM classification is combined with Quadratic Discriminant based classification to obtain a better tumor profile.

5. REFERENCES

- [1] M. Clark, et al., "Automatic tumor segmentation using knowledge-based techniques," *IEEE Transactions on Medical Imaging*, vol. 17, pp. 187-201, 1998.
- [2] M. R. Kaus, et al., "Automated Segmentation of MR Images of Brain Tumors," *Radiology*, vol. 218, pp. 586-591, 2001.
- [3] M. Schmidt, et al., "Segmenting Brain Tumors using Alignment-Based Features," presented at The Fourth International Conference on Machine Learning and Applications, Los Angeles, CA, 2005.
- [4] M. Prastawa, et al., "A brain tumor segmentation framework based on outlier detection," *Medical Image Analysis*, vol. 8, pp. 275-283, 2004.
- [5] S. Mori, et al., "Brain white matter anatomy of tumor patients evaluated with diffusion tensor imaging," *Annals of Neurology*, vol. 51, pp. 377-380, 2002.
- [6] S. M. Smith, et al., "Advances in functional and structural MR image analysis and implementation as FSL," *Neuroimage*, vol. 23, pp. 208-219, 2004.
- [7] M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Medical Image Analysis*, vol. 5, pp. 143-156, 2001.
- [8] G. J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*. NY: Wiley, 2004.
- [9] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167, 1998.