Hierarchical Segmentation of Malignant Gliomas via Integrated Contextual Filter Response

Shishir Dube[1], Jason J Corso[2], Alan Yuille[3], Timothy F. Cloughesy[4], Suzie El-Saden[1], and Usha Sinha[1]

[1]Medical Imaging Informatics (MII) Group, University of California, Los Angeles
[2]Laboratory of Neuro Imaging, University of California, Los Angeles
[3]Department of Statistics, University of California, Los Angeles
[4]Department of Neurology, University of California, Los Angeles

Abstract

We present a novel methodology for the automated segmentation of Glioblastoma Multiforme tumors given only a high-resolution T1 post-contrast enhanced channel, which is routinely done in clinical MR acquisitions. The main contribution of the paper is the integration of contextual filter responses, to obtain a better class separation of abnormal and normal brain tissues, into the multilevel segmentation by weighted aggregation (SWA) algorithm. The SWA algorithm uses neighboring voxel intensities to form an affinity between the respective voxels. The affinities are then recursively computed for all the voxel pairs in the given image and a series of cuts are made to produce segments that contain voxels with similar intensity properties. SWA provides a fast method of partitioning the image, but does not produce segments with meaning. Thus, a contextual filter response component was integrated to label the aggregates as tumor or non-tumor. The contextual filter responses were computed via texture filter responses based on the gray level co-occurrence matrix (GLCM) method. The GLCM results in texture features that are used to quantify the visual appearance of the tumor versus normal tissue. Our results indicate the benefit of incorporating contextual features and applying non-linear classification methods to segment and classify the complex case of grade 4 tumors.

Keywords: Segmentation, Classification, Pattern Recognition, Texture

Description of Purpose

Automated segmentation of brain tumors from MR images has been an extremely challenging problem due to the variable presentation of tumors, especially the malignant tumor, glioblastoma multiforme (GBM). GBMs present with varying extent of contrast enhancement, with or without significant edema, and further, with or without necrosis. The ability to automatically segment tumor and its sub-regions and then extract quantitative morphological features is of great significance to the management of subjects with brain tumors [1]. Several groups have addressed the problem of tumor segmentation from MR images. Fletcher-Heath et al. take a fuzzy clustering approach to the segmentation followed by 3D connected components to build the tumor shape [2]. Liu et al. take a similar fuzzy clustering approach in an interactive segmentation system that is shown to give an accurate estimate of tumor volume [3]. Kaus et al. use the adaptive template-moderated classification algorithm to segment the MR image into five different tissue classes for low grade gliomas: background, skin, brain, ventricles, and tumor [4]. Prastawa et al. present a detection/segmentation algorithm based on learning voxel-intensity distributions for normal brain matter and detecting outlier voxels, which are considered tumor [5]. Some groups have formulated segmentation techniques for single-channel images. Rajapakse et al. applied a statistical model to segment major tissue classes in a single T1-weighted SPGR sequence [6]. Despite the efforts of these groups, the problem of robust tumor segmentation remains a challenging one due to the variable appearance of these tumors as well as the variability in image quality of routine MR imaging studies.

Corso et al had earlier investigated a newly proposed 3D algorithm, multilevel segmentation by weighted aggregation (SWA) with integrated Bayesian model classification to segment GBM tumors from MR images. The algorithm worked on multi-channel image volumes, including T1-, T2-, FLAIR- weighted sequences, so that a unique characterization of each tissue type (brain, tumor, and edema) could be obtained. The results were promising, confirming the effectiveness of incorporating class modeling in the SWA process to adaptively segment abnormalities. The study obtained a Jaccard score of 0.85 for tumor based on multi-channel image sets with T1 pre-, post-contrast and FLAIR sequences. However,

Medical Imaging 2008: Image Processing, edited by Joseph M. Reinhardt, Josien P. W. Pluim,
Proc. of SPIE Vol. 6914, 69143Y, (2008)
1605-7422/08/$18 · doi: 10.1117/12.770769

this algorithm was based on images acquired from 3 or more different series which were acquired at low resolutions, with slice gaps and images between different series were often not aligned to each other. This last problem was circumvented by re-aligning images to a reference series but the resliced image series suffered from lower resolution along the slice axis as well as slice gaps. Imaging data often includes one high resolution image series (routinely this is the post contrast T1 weighted image volume) [7].

We now propose a modified SWA algorithm to segment tumor from brain tissue using single channel high resolution images. The challenge to be addressed is the intensity overlap between different tissue types in the single channel that confound the aggregation and classification. Following the original implementation of SWA, the intensity image is used to create the aggregates in the SWA algorithm. SWA, however, does not generate the final segmentation. This final step is accomplished, in our work, by classifying aggregates in three levels of the hierarchy. The novel aspect of this paper is the inclusion of contextual information through texture descriptors as features additional to intensity to classify aggregates as brain or tumor. Both linear and non-linear classifier performance was evaluated using precision and recall indices (with reference based on manual contours by expert). A majority voting mechanism based on the assignment from each of the three levels was used for the final labeling (brain or tumor) of a voxel.

Methods

The first step in our methodology was pre-processing the brain tumor training data. All image studies consisted of a T1 post-contrast enhanced SPGR sequence that had dimensions of 256x256x120 with an isotropic voxel resolution of 1 mm^3. The following steps were performed to reduce noise, remove non-brain matter, and obtain consistent image intensities across all subjects for the given MR channel: 1. Denoising – fMRI Software Library (FSL) nonlinear filter [8]; 2. Skull Stripping – FSL Brain Extraction Tool [8]; 3. Intensity Standardization – Performed via a normal hemisphere-based histogram matching method modified from [12]. For a given patient, the intensity histogram is extracted from the hemisphere that has the least presence of tumor (to prevent skewing of the intensity distribution from abnormalities). The extracted intensity histogram is then matched to a template histogram via a second-order polynomial transformation, and the transformation parameters are applied to the entire patient brain image to obtain a standardized volume (Figure 1).

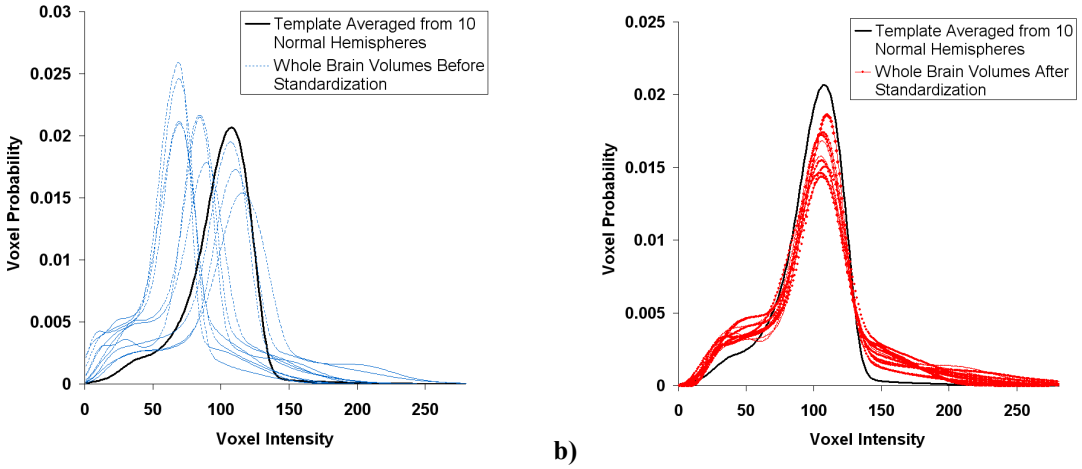

Fig. 1. a) T1CE histogram plots of $h_{template}$ and patient data before standardization b) T1CE histogram plots of $h_{template}$ and patient data after standardization

Sharon et al initially proposed SWA and the algorithm is detailed in the original paper as well as in the model based extension proposed by Corso et al [7,9]. We provide a brief overview here. Define a graph $G_t = (V, E)$ with the additional superscript indicating the level in the pyramid of graphs $G = \{G_t: t = 0,\ldots,T\}$. The finest layer in the graph $G_0 = (V_0,E_0)$ is created by the voxel lattice: each voxel i becomes a node $v \varepsilon V$ with a 6-neighbor connectivity, and node properties set according to the image, $s_v = I(i)$ where I(i) is the intensity at voxel i. The affinities, w_{uv}, are initialized

using $D(s_u,s_v;\theta) = \theta \mid s_u - s_v \mid_1$. Then SWA iteratively coarsens the graph by finding a set of representative nodes at each level in the hierarchy. This process continues until only one representative node remains. SWA provides a fast method for obtaining aggregates whose voxel members have strong affinity via the image intensity properties but it does not give the final segmentation. Further efforts are required to label the aggregates. Class assignment using intensity alone did not provide sufficient discrimination between brain and tumor. For example, enhancing blood vessels were often mis-classified as enhancing tumor. The addition of texture features significantly reduced such mis-classifications.

3D Texture volumes were calculated based on gray-level co-occurrence matrices (GLCM) obtained with a specified window size in the x-, y-, and z- directions. The GLCM can be calculated for different distance separation between voxel pairs (d), direction (θ), and neighborhood window size. Since GBM tumors have a highly variable nature, we chose a 3D neighborhood window (7 voxels in each dimension) with d = 1 and 26 different angular directions. The orientation of GBM is variable, but normal brain structures have a fairly consistent orientation. Thus, we chose to keep all resulting texture maps for the 26 different directions separate which resulted in 104 contextual channels. We chose the 3D neighborhood size so that it would be large enough to obtain meaningful information of the tumor lesion, but small enough such that smaller tumor lesions are also captured. To prevent sparse co-occurrence matrices and improve processing time, the MR image intensity range was reduced to 52 gray levels. The texture features that were calculated were: angular second momentum, contrast, entropy, and homogeneity [10].

Two different methods for classification were used for comparison: Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM). The two methods were chosen to evaluate the effectiveness of linear and non-linear classifiers. For a given aggregate at a specific level, the average intensity as well as the average texture feature from each contextual filter response was computed resulting in a feature vector with 105 elements. Three training data sets were obtained from levels 4 thru 6 since coarser levels result in aggregates that mix intensities from tumor and normal regions, and aggregates from finer levels are harder to distinguish between normal and abnormal. The two classes represented in the training set for classification were normal brain matter and tumor regions. For SVM classification, we decided to use a radial basis function kernel. To obtain the optimal parameter values for penalty (C_{level}) and gamma (γ_{level}) for each level, a 10-fold cross validation was performed on the corresponding training data with varying values of C_{level} and γ_{level} till a high level of training accuracy was reached. Finally, the best values for C_{level} and γ_{level} for each level were used to train on the entire corresponding training data set.

The ground truth data (brain and tumor regions) was obtained for 7 patient volumes; 5 of these were used in training the classifier and the other 2 for testing. For images in the training set, the average features (intensity and texture) for aggregates at each level were used to train the classifier. A discriminant classifier is then applied on the training set to determine the function that assigns a class label to a test aggregate sample. For a given level, the corresponding classifier determines a class for a particular aggregate as tumor or normal, and assigns the determined class to all voxels that are members of the test aggregate (Figure 3). To extract tumor from the three levels, we perform a majority vote of the class assignments of a given voxel at each level of the pyramid. The voxel is assigned a final class that has the greatest frequency of the set of class assignments within the three levels of the pyramid (Figure 4).

Results and Discussion

The intensity and texture images for a representative slice are shown in Figure 2 along with the truth data obtained from manual segmentation. A visual examination of these images shows that contextual information provides a good discrimination between tumor and brain. Further, the necrosis and cerebrospinal fluid have overlapping intensities (Fig. 2a), which are better differentiated in some of the contextual images (e.g., the homogeneity index image in Fig. 2e). Thus, the contextual information obtained from the texture filter responses is vital because the resultant features can more easily discriminate between tumor and normal brain matter regions than by intensity alone. It is clear that aggregates are larger (and fewer) as one goes up in the hierarchy. However, the choice of these levels for classification was made empirically and coarser levels were avoided where aggregates of brain and tumor may mix. The precision and recall scores were computed for the tumor segmentation and classification with both classifiers on the five training sets and the two test sets and are summarized in Table 1.

The use of intensity as a feature in classification required normalization across imaging studies. Figure 1a shows that the raw images had a wide distribution of intensities that would have confounded intensity-based classification. The

histogram based intensity normalization allowed for reproducible intensity ranges for each image study (Fig. 1b). However, even standardized intensity distributions showed overlapping histograms for different tissue types. Thus, classification based on intensity values alone reflected this overlap in low precision and recall values for both classifiers (Table 1). Clearly, overlapping clusters with one feature alone are better separated using two features. This discriminatory power of the combined features is seen in the higher recall and precision indices using all features (Table 1).

Earlier work by Akselrod-Ballin applying SWA to segmentation of multiple sclerosis lesions included a decision tree based top down classification scheme on randomly selected aggregates at three levels (small, intermediate, large) [11]. In contrast, we chose to classify aggregates at only relevant levels where the tumor region was almost completely segmented (unlike at finer levels) and yet not merged with the brain tissue (unlike at coarser levels). The limitation of our method is the empirical choice of the relevant levels. The decision tree based classification of Akselrod-Ballin is certainly more sophisticated than the one proposed here but it applies Fishers Linear Discriminant to partition the aggregates at each node. Our experience comparing SVM to LDA shows that tissue classes may be optimally separable using non-linear classifiers. It is likely that classification incorporating SVM with the decision tree approach may produce higher accuracies of segmentation.

An examination of the performance of the two classifiers showed the superior performance of the SVM in precision indices. However, its recall was poor in a few cases and corresponded to necrotic regions being mis-classified as brain. In most of the segmentations, the LDA classifier mis-classified enhancing blood vessels as tumors, which lead to poor precision values.

There are several avenues for future work. The choice of the texture calculation at a window of 7 voxels, d=1 was determined empirically. We propose to perform an exhaustive search for the optimal texture features using different windows and distance values. Smaller window sizes may reduce the blur in the texture images using the 7x7x7 window used in the current implementation. Furthermore, in the current implementation the texture features were used only in the classification rather than being integrated in the aggregation process itself. This allows a multi-channel version using the intensity and texture images. Other options include computation of texture over each aggregate, rather than pre-computed values using a fixed window size. Clearly, it is important to identify the most discriminant features before using them either for classification or directly in the SWA process.

Conclusions

In this paper, we presented a novel application of contextual filter response coupled with SWA to classify tumor regions using a single channel volumetric T1 post contrast enhanced sequence. The textual features provided discrimination not possible on intensity alone. The features coupled with the SVM classifier shows promise for robustly locating tumor regions from a single channel high-resolution T1 post-contrast image.

References

1. Patel, M.R., Tse, V.: Diagnosis and staging of brain tumors. Seminars in Roentgenology 39(3) (2004) 347-360
2. Fletcher-Heath, L.M., Hall, L.O., Goldgof, D.B., Reed Murtagh, F.: Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. Artificial Intelligence in Medicine 21 (2001) 43–63
3. Liu, J., Udupa, J., Odhner, D., Hackney, D., Moonis, G.: A system for brain tumor volume estimation via MR imaging and fuzzy connectedness. Computerized Medical Imaging and Graphics 29(1) (2005) 21–34
4. Kaus, M., Warfield, S., Nabavi, A., Black, P.M., Jolesz, F.A., Kikinis, R.: Automated segmentation of MR images of brain tumors. Radiology 218 (2001) 586–591
5. Prastawa, M., Bullitt, E., Ho, S., Gerig, G.: A brain tumor segmentation framework based on outlier detection. Medical Image Analysis Journal, Special issue on MICCAI 8(3) (2004) 275–283
6. Rajapakse JC., Giedd JN., Rapoport JL.: Statistical Approach to Segmentation of Single-Channel Cerebral MR Images. IEEE Transactions on Medical Imaging 16(2) (1997) 176-186
7. Corso JJ., Sharon E., Yuille A.: Multilevel Segmentation and Integrated Bayesian Model Classification with an Application to Brain Tumor Segmentation. In: Proceedings of MICCAI 9(2) (2006) 790-798
8. Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., Luca, M.D., Drobnjak, I., Flitney, D.E., Niazy, R., Saunders, J., Vickers, J., Zhang, Y., Stefano, N.D., Brady, J.M., Matthews, P.M.: Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 23(S1) (2004) 208–219

9. Sharon, E., Brandt, A., Basri, R.: Fast multiscale image segmentation. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Volume I. (2000) 70–77
10. Mahmoud-Ghoneim D, Toussaint G, Constans JM, de Certaines JD.: Three dimensional texture analysis in MRI: a preliminary evaluation in gliomas. Magnetic Resonance Imaging 21(9) 2003 983-987
11. Akselrod-Ballin, A., Galun, M., Gomori, M.J., Filippi, M., Valsasina, P., Basri,R., Brandt, A.: Integrated segmentation and classification approach applied to multiple sclerosis analysis. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. (2006)
12. Nyul LG, Udupa JK, Zhang X, New variants of a method of MRI scale standardization, IEEE Trans Med Imaging. 2000 Feb;19(2):143-50

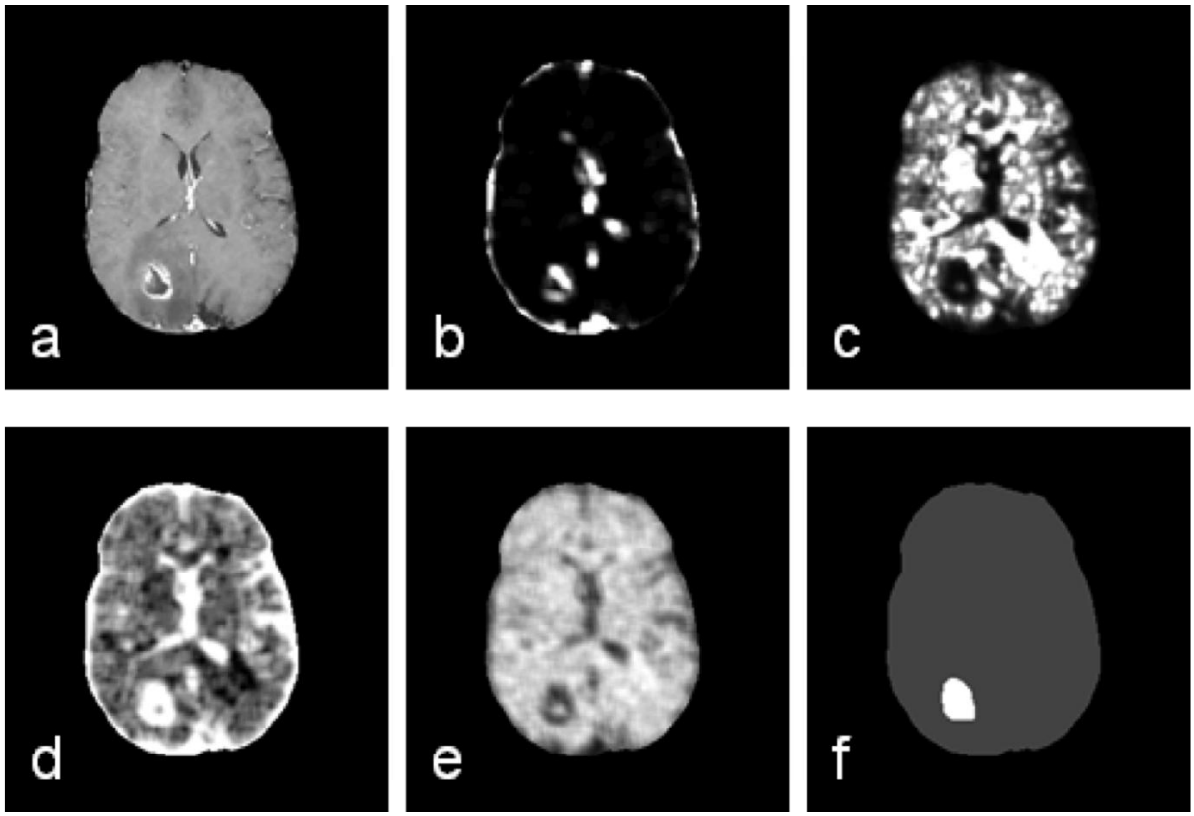

Fig. 2. The T1 post-contrast enhanced intensity image (a) for a given patient as well as the respective contextual filter response images and ground truth at the corresponding slice: (b) contrast, (c) energy, (d) entropy, (e) homogeneity, (f) ground truth.

	Intensity				All Features			
	LDA Precision	LDA Recall	SVM Precision	SVM Recall	LDA Precision	LDA Recall	SVM Precision	SVM Recall
Train 1	0.2	0.7	0.39	0.59	0.39	0.7	0.85	0.36
Train 2	0.42	0.4	0.93	0.29	0.56	0.67	0.873	0.612
Train 3	0.66	0.721	0.77	0.49	0.35	0.05	0.57	0.83
Train 4	0.673	0.808	0.87	0.53	0.661	0.86	0.92	0.83
Train 5	0.486	0.603	0.75	0.39	0.526	0.661	0.84	0.59
Test 1	0.45	0.27	0.74	0.07	0.246	0.528	0.77	0.23
Test 2	0.3	0.54	0.39	0.2	0.342	0.749	0.95	0.45

Table 1. Precision/Recall values for the tumor segmentation for five training and two test cases with image intensity as the only feature and the inclusion of intensity and contextual features.

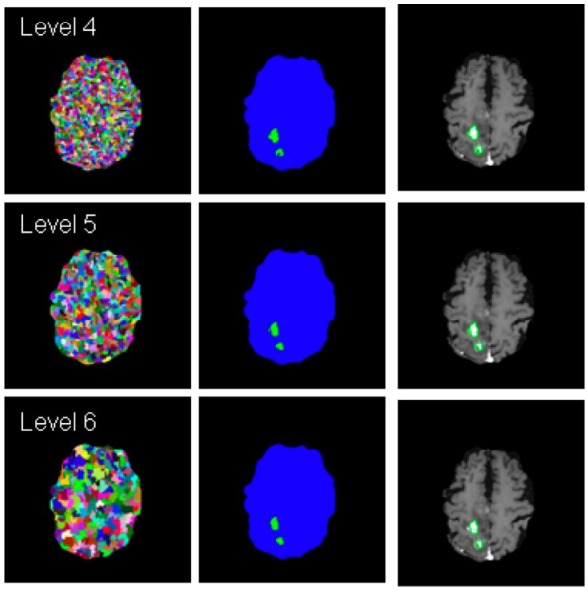

Fig. 3. Each row shows the level in the pyramid hierarchy, the segmented tumor (green) and brain (blue), and the tumor boundaries outlined on the corresponding intensity image slice.

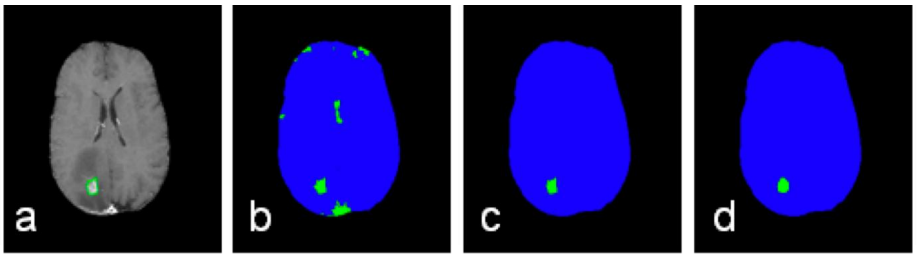

Fig. 4. Images showing final segmentation of tumor (green) and brain (blue) tissues. (a) A T1 post contrast enhanced image with an outline of the tumor segmentation resulting from SVM classification on all features, (b) segmentation resulting from LDA classification on all features, (c) segmentation resulting from SVM classification on all features, (d) ground truth