MEDINFO 2007
HISA Ltd Australia
© 2007 The authors. All rights reserved.

Content Based Image Retrieval of Glioblastoma Multiforme

Shishir Dube[a], Suzie El-Saden, Timothy F. Cloughesy[b], Usha Sinha[a]

[a]*Department of Biomedical Engineering, University of California at Los Angeles,CA,USA*
[b]*Department of Neurology, University of California at Los Angeles,CA,USA*

Abstract

We propose a Content Based Image Retrieval system, for patients diagnosed with Glioblastoma Multiforme that will predict of time to survival and allow a neuroradiologist to modify treatment procedures based on quantified features explicitly extracted from segmented regions of the tumor. Our proposed system has two components: a preprocessing scheme to improve the image quality and provide consistency and a multivariate linear model. The multivariate linear model applied to the training data had a correlation coefficient of 0.848, which indicated a strong association of the selected features to time to survival. Future work will involve expanding the training set and incorporating additional features not explicitly extracted from the segmented tumor regions.

Keywords:

content based image retrieval, imaging informatics, image data mining, image processing, disease modeling

Introduction

Medical image databases have grown immensely in the past few years. Imaging studies, such as magnetic resonance (MR) and computed tomography (CT) result in a large volume of data. The ability to query large image databases can form the basis of a decision support system. However, a large number of existing image databases are indexed by text annotations that routinely contain patient demographic or study information. Manual methods to extend this with more comprehensive text annotation that capture image content are both tedious and time-consuming. We propose a CBIR system for GBM tumor cases for the purposes of visualizing the variety of different disease presentations that can exist given similar features as well as prediction of time to survival given a query volume (for the purposes of treatment modification). Our proposed CBIR system includes a fast and automated segmentation component with the ability to allow a radiologist to modify the pertinent regions segmented, as well as a multivariate linear regression model predicting time to survival based on particular features (extracted only from the T1 post contrast enhanced studies) of the tumor region.

Materials and methods

Two important components are required in order to have an accurate and efficient CBIR system of brain tumor images: I) preprocessing of images to reduce variability resulting from differing MR scanner conditions; II) an accurate model relating time to survival to extracted quantified features. Our proposed methods for the three areas are discussed below.

Preprocessing scheme

The following steps were performed to improve SNR, register the two image channels of a given subject, and to obtain consistent image intensities across all subjects for a given channel: Denoising –FSL nonlinear filter [3], Registration – For image alignment [3], Intensity Standardization – Performed via a histogram matching method (Figure 1) [4].

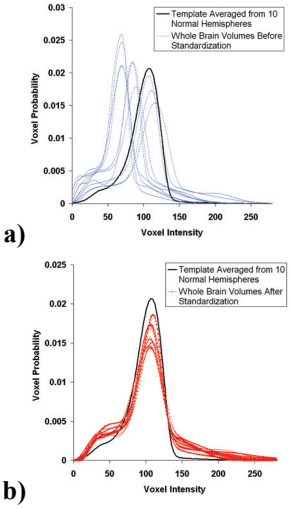

Figure 1 - a) Histogram plots of $h_{template}$ and patient data before standardization b) Histogram plots of $h_{template}$ and patient data after standardization

Quantitative feature modeling

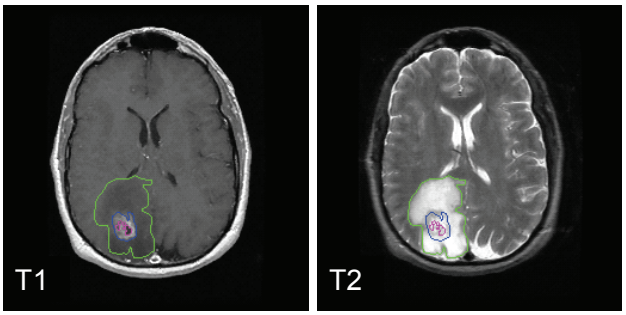

Figure 2 - The left and right images are the T1 post contrast and respective T2 with the labeled regions that were segmented for quantitative analysis. The edema boundary is outlined by green, contrast enhancement by blue, and necrosis by magenta

Imaging data from 34 subjects with GBM were included in the analysis for training purposes, and an additional 4 patients' images were kept aside for querying purposes. The quantitative features were extracted from only the T1 post-contrast study, but the T2 sequence was used to accurately delineate the edema region (Figure 2). Table 1 is a list of the shape, texture and intensity histogram features extracted for each tumor sub-region. A stepwise multivariate linear regression model was applied to the quantitative data to determine the quantitative variable (or combination of variables) that best predict time to survival.

Table 1 - List of the different features extracted for each sub-region. The features consisted of texture parameters derived from 3D co-occurrence matrices, 3D shape moments, and histogram parameters

Category	Features
3D Texture via Co-occurrence Matrix	Correlation Contrast Homogeneity Energy
3D Shape Moments	Order 0 Invariant Moment (Volume) Order 1 Invariant Moments Order 2 Invariant Moments Order 3 Invariant Moments
Histogram	Mean Variance Skewness Kurtosis

Results

Multivariate Linear Regression Model

Table 2 shows the results of the stepwise multivariate linear regression model.

Table 2 - Results of the stepwise multivariate linear regression model applied to the quantitative imaging data

Model	Variables	Coefficients	Adjusted R^2
1	Constant	329.54	0.648
	3rd order SM of CE	4.37E-06	
2	Constant	326.154	0.739
	3rd order SM of CE	4.27E-06	
	2nd order SM of Edema	-8.40E-06	
3	Constant	404.316	0.783
	3rd order SM of CE	4.51E-06	
	2nd order SM of Edema	-6.50E-06	
	Correlation of Edema	788.71	
4	Constant	406.22	0.814
	3rd order SM of CE	4.55E-06	
	2nd order SM of Edema	-5.50E-06	
	Correlation of Edema	771.308	
	3rd order SM of Edema	6.34E-07	
5	**Constant**	**214.446**	**0.848**
	3rd order SM of CE	**4.61E-06**	
	2nd order SM of Edema	**-3.80E-06**	
	Correlation of Edema	**747.286**	
	3rd order SM of Edema	**6.85E-07**	
	Kurtosis of CE	**64.09**	

Table 2 shows the results of the stepwise multivariate linear regression model. Each model corresponds to the number of variables used to predict survival and the respective correlation coefficient. For a given model n (where n represents the number of variables), the variables (along with their respective weighting coefficients) that gave the best correlation with time to survival are listed.

Discussion

As can be seen in Table 2, the shape and texture of edema show strong correlations to time to survival. Edema presumably follows the white matter tracts, and a higher infiltration of the tracts may be captured in the shape indices as structures of high morphological anisotropy. The correlation texture feature is related to the inhomogeneity of intensity in the edema sub-region. Infiltration of tumor cells into edema could potentially result in inhomogeneous intensities of the edema sub-regions. Greater infiltration into the surrounding peri-tumoral edema has a poor prognosis and this could be the basis for a correlation between the edema texture feature and time to survival. The correlation of the shape of contrast enhancing tumor region to survival time may also have its origin in the infiltrative nature of tumor. Kurtosis of the intensity histogram primarily reflects asymmetry and the presence of a long tail in the high intensity end of the histogram. This heterogeneous intensity profile may arise from some voxels

showing a high uptake of the contrast agent, which in turn may reflect tumor aggressiveness.

Future work will involve expanding the multivariate model training set to 100 patients and including features not derived explicitly from the segmented volumes (i.e. age, presence of multifocality) that have been shown to have some correlation with survival.

References

[1] Sinha U, Kangarloo H. Principal Component Analysis for content-based image retrieval. *Radiographics* 2002; 22:1271-89.

[2] Pope WB, Sayre J, Perlina A, Villablanca JP, Mischel PS, Cloughesy TF. MR Imaging Correlates of Survival in Patients with High-Grade Gliomas. *AJNR* 2005; 26:2466-2474.

[3] Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens T, Johansen-Berg H, Bannister PR, Luca MD, Drobnjak I, Flitney DE, Niazy R, Saunders J, Vickers J, Zhang Y, Stefano N, Brady JM, Matthews PM. Advances in functional and structural mr image analysis and implementation as fsl. *NeuroImage* 2004; 23:208–219.

[4] Madabushi A, Udupa JK. New methods of MR image intensity standardization via generalized scale. *Med Phys* 2006; 33: 3426-34.