RADIOMICS TEXTURE FEATURE EXTRACTION FOR CHARACTERIZING GBM PHENOTYPES USING GLCM

Ahmad Chaddad[a], Pascal O. Zinn[b] and Rivka R. Colen[a]

[a]Department of Diagnostic Radiology, MD Anderson Cancer Center, University of Texas, USA
[b]Department of Neurosurgery, Baylor College of Medicine, USA

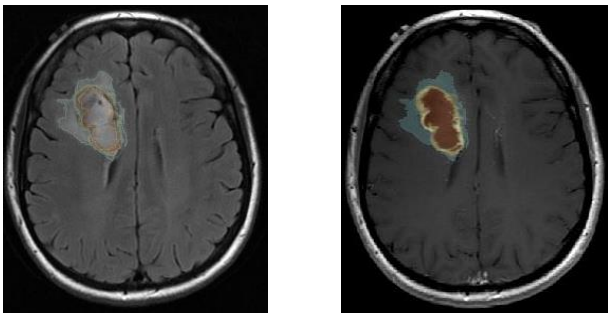
ABSTRACT

Glioblastoma (GBM) is a markedly heterogeneous brain tumor and is composed of three main volumetric phenotypes, namely, necrosis, active tumor and edema, identifiable on magnetic resonance imaging (MRI). This paper assesses the usefulness of the GBM features detection by using semi-automatic segmentation and texture feature extracted from gray level co-occurrence matrix (GLCM). Feature vectors are then used for predicting GBM phenotypes based on nearest neighbors (NN) classifier. Simulation results for 22 patients show an accuracy of 75.58% for distinguishing GBM phenotypes based on the texture feature selection using the decision trees model. Preliminary texture analysis demonstrated that the texture feature based on the GLCM is promising to distinguish GBM phenotypes.

Index Terms— Glioblastoma, Texture, GLCM, MRI

1. INTRODUCTION

Medical image analysis widely uses texture feature extraction for getting more characters and information [1-2]. Many experiments of texture analysis have been applied in medical imaging such as: radiographics [2], ultrasound [3], computed tomography [4], PET [5], MRI [6-7]. This study focuses on the GBM phenotypes which are represented by the three volumetric namely, necrosis (vN), active tumor (vAT) and edema (vE). GBM tumor is a heterogeneous tissue with intratumoral spatial variation in cellularity, and its heterogeneity is a pattern feature of malignancy which is represented by areas of high cell density. Heterogeneity in this context can be quantified by several texture feature types derived from Gaussian mixture model [7], Histogram [8], and Wavelets [9]. Texture analysis using dynamic contrast-enhanced MRI (DCE-MRI) was done to discriminate GBM from malignant glioneuronal tumors (MGNT) [6]. It used application of both types; first-order and second-order statistics based on GLCM and run-length matrix (RLM) methods. Its results were promising, for example: run-length non-uniformity, gray-level non-uniformity, angular second moment, and entropy to the findings from DCE-MRI had 100 % negative predictive value and sensitivity in discriminating MGNT from GBM [6]. Another aspect of texture analysis was employed by using Gabor transform for combining conventional MRI and perfusion MRI. This technique was successfully done to discriminate metastases from gliomas and also, high grade from low grade glioma [10]. In this paper, we propose a new approach of GBM phenotype analysis by using a texture feature extracted from GLCM and then selected by decision tree model. Our method considers co-occurrence pair pixels which detect the heterogeneity variation for each GBM phenotype. This technique can characterize each phenotype by a specific set of features, and subsequently can be used for automatically predicting GBM features. It might be able to provide a more accurate assessment of the patient prognosis and underlying genomic composition. Note that GBM phenotypes are depicted by using registration model applied on T1-weighted (T1-WI) and its corresponding fluid attenuated inversion recovery (FLAIR sequence (Fig.1).



(a) (b)

Fig. 1. Example of GBM phenotypes, (a) Segmentation based on the registered using MR T1-WI and FLAIR sequence, (b) necrosis, active tumor and edema are labeled by the red, yellow and cyan colors respectively

2. MATERIALS AND METHODS

Schematic representation of a robust method for the GBM features detection by using raw data from MRI is shown in Fig. 2. Raw images have dimensions of 256×256 pixels. The images are converted into grayscale before further processing. A dataset of 22 patients with grade 4 glioma types "GBM", image has been analyzed by using rigid

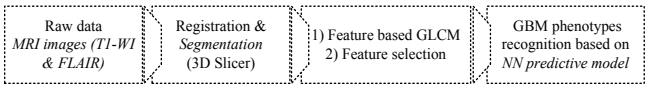


Fig. 2. Schematic diagram of GBM phenotypes discrimination. Four main steps in designing for robust *GBM* phenotypes detection are shown.

registration of T1-WI and its corresponding FLAIR sequence. The age of the patients in the dataset ranges from 30 to 70 years. Semi-automatic segmentation for GBM phenotypes are employed by using 3D Slicer tools. Features based on GLCM and feature selection are applied on 3 GBM phenotypes, and NN predictive model is used for GBM features discrimination.

2.1. Registration and Segmentation

3D Slicer tool was used to register T1-WI and its corresponding FLAIR sequence. The two corresponding scans or images were rigidly aligned to each other to make an accurate registration (Fig. 3c). Additionally, we segmented the phenotype in each axial slice by certified neuroradiologists who determined the three phenotypes using 3D Slicer tool [11]. The necrosis and active tumor are determined using T1-WI (Fig. 3d,e), and edema/ invasion by using FLAIR sequence (Fig. 3f). An expertise neuroradiologist validated all these segmentations.

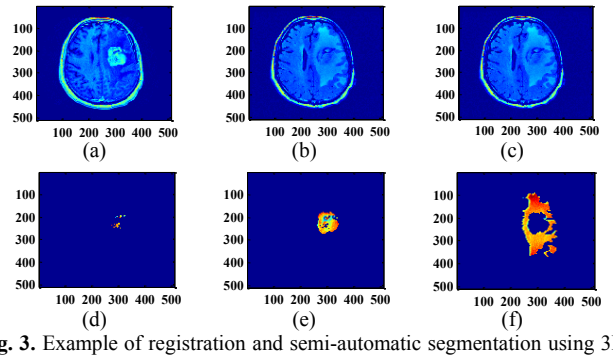


Fig. 3. Example of registration and semi-automatic segmentation using 3D Slicer,(a) T1-WI image, (b) corresponding FLAIR sequence, (c) corresponding registration, (d) necrosis, (e) active tumor, and (f) edema/invasion.

2.2. Texture Feature Based GLCM

Texture features were extracted from GLCM of each GBM phenotype by using Matlab 2013. GLCM was employed to perform two-dimensional texture analysis. The texture features analysis was performed on raw data of registered images as follows. First, for each phenotype image, histogram equalization was applied to 256 levels, accounting only for the segmented phenotype pixels. Second, GLCM was calculated based on specific offset and phase. Note that we employed 8 different phases (0°, 45°, 90°, 135°, 180°, 225°, 270° and 315°) and 3 offsets (1, 2 and 3 pixels). Moreover, GLCM size consists of 256×256 elements and each element value refers to the number of times that the pair pixels are quantified by specific offset and phase. There are 24 GLCM matrices of size 256×256 for each GBM phenotype image. The texture features were calculated from each GLCM and the values gained for each feature at the different offset and phases used for predicting GBM phenotype [12]. 528 features were generated based on 24 GLCM quantified by 22 texture features. The use of high number of texture features can be a cause for information redundancy. To resolve this problem, we used decision tree model in order to identify a subset of the features that is most relevant for necrosis, active tumor and edema/invasion discrimination [13].

2.3. NN Predictive Model and Performance Metrics

We implemented a simple supervised learning algorithm which is commonly used in learning and classification. In NN model, a phenotype sample is classified based on the distance of its features from those of its neighbors, with the sample being assigned to the class most common among its K (K is the number of nearest neighbors samples) nearest neighbors. The neighbors are taken from a set of samples, called the training set, for which the correct classification is known. In this study, K (K> 86, 86 is the number of each GBM phenotype samples) is chosen to be odd when the number of classes is 3 to resolve any ties. A higher K increases the classification accuracy but at the expense of computational time. To validate our results, we calculated 4 metrics, namely, accuracy, sensitivity, specificity and negative predictive value (NPV), to provide a quantitative assessment of the predictive model. These metric are defined as follows.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

where the true positive (*TP*) and the true negative (*TN*) are the number of correctly classified positive and negatives classes. The false positive (*FP*) and false negative (*FN*) are those samples which are incorrectly classified.

Sensitivity: It evaluates the capability of a classifier to recognize the positive class patterns. It can be expressed according to

$$Sensitivity = \frac{TP}{TP+FN} \quad (2)$$

Specificity: It evaluates the capability of a classifier to recognize the negative class patterns. It can be expressed by the following equation

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

NPV: It is the probability that samples with a negative screening test are truly negative. It can be calculated using the following expression:

$$NPV = \frac{TN}{FN+TN} \quad (4)$$

where *TN* and *FN* are true negative and false negative, respectively.

Similarly, receiver operating characteristic (ROC) curve is employed which provides the true positive versus (vs) false positive rates which is associated with area under the curve (AUC) [14]. Finally, the metrics of the texture features (1, 2 and 3 pixels), cumulative features, full feature set and feature selection were calculated and reported.

3. RESULTS AND DISCUSSION

To assess the value of texture features to identify phenotypes differences in GBM, we performed an integrated analysis assessing prognostic performance and association with phenotypes data sets. First, we defined 528 quantitative texture features describing phenotype characteristics by texture feature extraction from GLCM. To investigate the effect of texture feature in predicting GBM phenotype we applied a comparative study between all the 528 texture features. Performance metrics of auto-correlation (offset=2, 0° and 180°) and correlation (offset=1, 135°) showed a maximum classifier accuracy of 66.28 and 65.5% respectively (Fig. 4a). Additionally, cumulative texture feature (CTF) showed an acceptable accuracy of ~71.00 % by using full feature set with stability of sensitivity (~73.26%), specificity (~79.07%) and NPV (~85.53%) metrics (Fig. 4b). The best accuracy (~72.48%) is represented by using the first 248 CTF. However, the best sensitivity and NPV are 76.74% and 87.01% respectively which are provided by using first 29 texture features. Maximum specificity with 84.88% is obtained by using first 6 texture features. Variance of performance metrics function with texture feature number is more stable when we employed range of 439-528 texture features (Fig. 4b). Area under curve (AUC) of (*vE* versus *vN)* is greater than (*vAT* versus *vN*), and (*vAT* versus *vE*) of 90.67, 86.88 and 75.54% respectively (Fig. 4c). Then, we performed a specific analysis assessing prognostic performance associated with specific offset and phases of GLCM. First we defined 3 groups of quantitative image features describing GBM phenotypes: each group is represented by 8 phases, and corresponding offset pixel of GLCM matrix (176 features), namely, offset 1, 2 and 3 respectively (Table 1). Additionally, we considered two other groups, namely, 248 CTF and Full feature set. Performance metrics showed maximum classifier accuracy of 70.93% by using full set features (528 texture features), and 70.16 % accuracy based on 2 offsets and 8 phases (176 texture features), and these were more than the accuracy using 1 and 3 offsets namely 67.05 and 68.99% respectively. Based on the representation of the feature groups (Table 1), we observed that the full set features are the most accurate to characterize the GBM phenotypes. However, the best accuracy of 73.20% is still represented by using first 248 CTF without specific phase and offset considered.

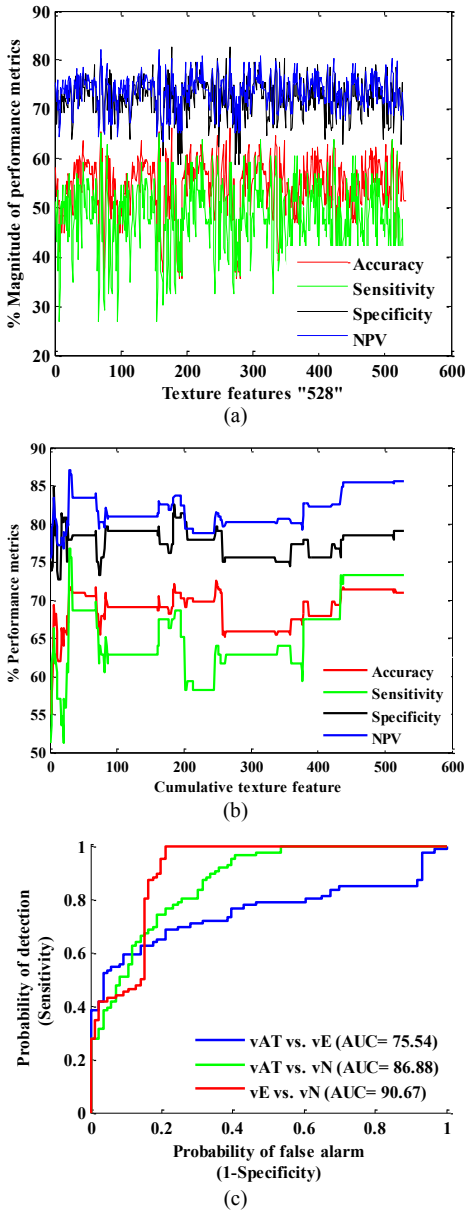


Fig. 4. Effect of texture features for GBM phenotypes discrimination, (a) analysis of performance metrics for each texture feature, (b) cumulative texture features, (c) ROC curves associated with AUC.

Additionally, highest accuracy of 75.58 % is showed by using 8 texture features obtained from decision tree which is a non-parametric approach which does not require any prior assumptions about the probability distributions of the various features [15]. These 8 texture features (C_1, SA, E_1, CP, AV, CA_1, SE_2 and C_2) were selected based on decision trees from feature collected offset (1, 2, and 3 pixels), full feature set and 248 CTF (Fig. 5). This approach can have a large impact of noninvasive imaging in a large number of GBM patients and the automated texture feature extraction. The preliminary result of this work could motivate further research of image-based quantitative texture features.

Table 1. Effects of texture features extracted from GLCM with 8 phases

Feature collected	Accuracy	Sensitivity	Specificity	NPV
Offset 1 pixel	68.99	67.44	77.33	82.61
Offset 2 pixels	70.16	63.95	80.81	81.76
Offset 3 pixels	67.05	65.12	77.33	81.60
Full set features	70.93	73.26	79.07	85.53
248 CTF	72.48	63.95	79.65	81.55
Feature selection	**75.58**	**63.95**	**90.69**	**83.42**

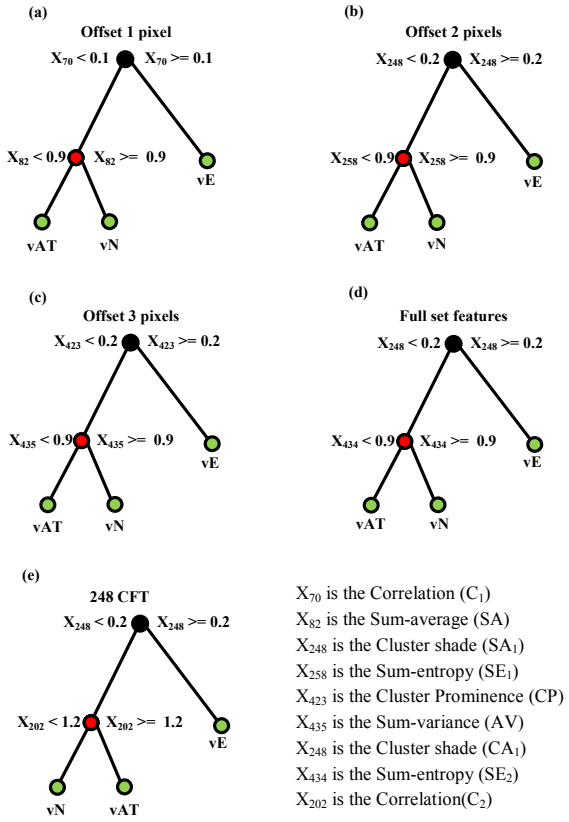


Fig. 5. Exploratory decision tree grown using 5 texture feature groups, and the features selection in each group are:, (a) C_1 and SA, (b) SA_1 and SE_1, (c) CP and AV, (d) CA_1 and SE_2, and (e) SA_1 and C_2.

The value of these quantitative features in detecting treatment response and subsequently patient outcome are topics of further research.

4. CONCLUSION

In this paper, we implemented a new approach for GBM phenotype discrimination. Thus, we presented rigid registration based on T1-WI and its corresponding FLAIR sequence, and semi-automatic segmentation of GBM phenotypes based on 3D Slicer tools have been applied. We investigated the quantization, offset, and phases factors of GLCM on MR GBM phenotypes imagery. Comparative study showed an advantage accuracy of CTF and feature selection technique. It is noticed that the proposed methodology has performed much better when we considered CTF and feature selection for GBM phenotypes discrimination. The performance of our approach can be further improved by extensive the number of GBM patients.

5. REFERENCES

[1] A. Chaddad, et al., "Extracted Haralick's Texture Features and Morphological Parameters from Segmented Multispectrale Texture Bio-Images for Classification of Colon Cancer Cells," *WSEAS Transaction on Biology and Biomedicine*, vol. 8, pp. 39–50, 2011.

[2] C. A. Harlow and S. A. Eisenbeis, "The Analysis of Radiographic Images," *IEEE Transactions on Computers*, vol. C–22, no. 7, pp. 678–689, Jul. 1973.

[3] G. D. Tourassi, "Journey toward Computer-aided Diagnosis: Role of Image Texture Analysis," *Radiology*, vol. 213, no. 2, pp. 317–320, Nov. 1999.

[4] B. Ganeshan, et al., "Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: Preliminary evidence of an association with tumour metabolism, stage, and survival," *Clinical Radiology*, vol. 67, no. 2, pp. 157–164, Feb. 2012.

[5] F. Tixier, et al., "Intratumor Heterogeneity Characterized by Textural Features on Baseline 18F-FDG PET Images Predicts Response to Concomitant Radiochemotherapy in Esophageal Cancer," *J Nucl Med*, vol. 52, no. 3, pp. 369–378, Mar. 2011.

[6] P.-A. Eliat, et al., "Can Dynamic Contrast-Enhanced Magnetic Resonance Imaging Combined with Texture Analysis Differentiate Malignant Glioneuronal Tumors from Other Glioblastoma?," *Neurology Research International*, vol. 2012, p. e195176, Dec. 2011.

[7] A. Chaddad, P. O. Zinn, et R. R. Colen, " Brain tumor identification using Gaussian Mixture Model features and Decision Trees classifier," *IEEE 48th Annual Conference on Information Sciences and Systems*, pp.1‑4, 2014.

[8] A. Chaddad and R. R. Colen, "Statistical feature selection for enhanced detection of brain tumor," *SPIE 9217, Applications of Digital Image Processing XXXVII*, 92170V – 92170V – 8, 2014.

[9] S. Dua, U. R. Acharya, P. Chowriappa, et S. V. Sree, "Wavelet-based energy features for glaucomatous image classification," *IEEE Trans Inf Technol Biomed*, vol. 16, n° 1, p. 80‑87, 2012.

[10] E. I. Zacharaki, et al., "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme," *Magn. Reson. Med.*, vol. 62, no. 6, pp. 1609–1618, Dec. 2009.

[11] "3D Slicer." [Online]. Available: http://www.slicer.org/. [Accessed: 1-April-2014].

[12] A. Chaddad, P. O. Zinn, et R. R. Colen, "Quantitative Texture Analysis for Glioblastoma Phenotypes Discrimination," *IEEE CoDIT*, pp.605‑608, 2014.

[13] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar *Introduction to Data Mining*, Addison-Wesley, 2005.

[14] D. J. Hand and R. J. Till, "A Simple Generalisation of the Area Under the ROC Curve for Multiple Class Classification Problems," *Machine Learning*, vol. 45, no. 2, pp. 171–186, Nov. 2001.

[15] L. Breiman, *Classification and regression trees*. Chapman & Hall, 1984.