Research Article High-Throughput Quantification of Phenotype Heterogeneity Using Statistical Features

Ahmad Chaddad and Camel Tanougast

Laboratory of Conception, Optimization and Modelling of Systems, University of Lorraine, 7 rue Marconie, Metz, 57070 Lorraine, France

Correspondence should be addressed to Ahmad Chaddad; ahmad.chaddad@univ-lorraine.fr

Received 24 May 2015; Revised 28 September 2015; Accepted 1 October 2015

Academic Editor: Klaus Jung

Copyright © 2015 A. Chaddad and C. Tanougast. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Statistical features are widely used in radiology for tumor heterogeneity assessment using magnetic resonance (MR) imaging technique. In this paper, feature selection based on decision tree is examined to determine the relevant subset of glioblastoma (GBM) phenotypes in the statistical domain. To discriminate between active tumor (*v*AT) and edema/invasion (*vE*) phenotype, we selected the significant features using analysis of variance (ANOVA) with *p* value < 0.01. Then, we implemented the decision tree to define the optimal subset features of phenotype classifier. Naïve Bayes (NB), support vector machine (SVM), and decision tree (DT) classifier were considered to evaluate the performance of the feature based scheme in terms of its capability to discriminate *v*AT from *vE*. Whole nine features were statistically significant to classify the *v*AT from *vE* with *p* value < 0.01. Feature selection based on decision tree showed the best performance by the comparative study using full feature set. The feature selected showed that the two features Kurtosis and Skewness achieved a highest range value of 58.33–75.00% accuracy classifier and 73.88–92.50% AUC. This study demonstrated the ability of statistical features to provide a quantitative, individualized measurement of glioblastoma patient and assess the phenotype progression.

1. Introduction

Glioblastoma is the most common tumor and most aggressive primary brain malignancy in adults [1]. The inability to perform complete surgical tumor resection and poor drug delivery to the brain contribute notably to the limited treatment options. Despite all efforts, the average of patient survival with GBM currently is thereabouts 14.6 months [2]. GBM consists of active tumor, peritumoral edema, and necrosis parts, as designated by the combination of T1weighted (T1-WI) and Fluid-Attenuated Inversion Recovery (FLAIR) images in MRI. The active tumor region is described as the contrast-enhancing portion in T1-WI images and peritumoral edema is defined as the hyperintense region of FLAIR images, located outside the active area. The recent improvements in MRI technology using the integration of diffusion and perfusion weighted imaging have provided deeper insights into the pathological behavior of tumors [3, 4].

Until now, radiologists have used MR imaging for relatively gross disease detection. We hypothesize that radiomics with the availability techniques in image processing applied on the raw data derived from MRI can make radiological examinations more effective. In this way, automatic data computation could foster faster and effective readings of numerous types of images and classify them as normal or cancerous [5]. Such a system must have the ability to detect and extract the abnormal areas from their surroundings by automatic segmentation techniques such as multithresholding segmentation technique with the morphological image processing [6]. In terms of tumor heterogeneity, technical research has investigated this heterogeneity by quantifying its texture using numerous functions. For instance, feature extraction based on the gray level cooccurrence matrix (GLCM) with Haralick features is a popular technique used for texture analysis [7]. Then, GLCM computes the neighborhood correlations around pixels where the GLCM is calculated by the paired pixel in specific offset (distance) and phase (direction). Statistical (histogram) analysis has been established for the pixel intensity or the map of pixel orientations. In this context, the statistical approach used has been the texture-based approach [8–12]. Previous studies of GBM assessment have required registration of T1-WI and FLAIR for identifying the phenotypes, and each of the visible phenotypes is segmented manually by board certified radiologists.

In this paper, a novel approach for analyzing GBM phenotypes using FLAIR images only is introduced. Histogram based statistical features can offer a simple way to characterize GBM heterogeneity across the phenotypes, namely, *v*AT and *vE*. Reproducible quantifiable imaging features of GBM heterogeneity that explicitly examine links between the imaging findings and the underlying GBM phenotypes characteristics are identified. Introduced quantitative histogram features can discriminate phenotype heterogeneity from MRI images and thereby strengthen personalized medicine in GBM [12].

2. Materials and Methods

To prove the hypothesis, we focused on the optimal subset features from the statistical features which are derived from GBM tumors using active GBM portion with high intensity pixels vAT and peritumoral vE of GBM with middle intensity pixels. Two Gaussian distributions could clearly be observed in the histogram data of GBM (Figure 1). To assist automated recognition of the GBM phenotype based heterogeneity, histogram statistical features and classifier techniques were used for discriminating active tumor parts from edema parts in FLAIR images. Decision tree was considered to recognize the dominant statistical features which represented the foremost characteristic of GBM heterogeneity [13]. The proposed approach is presented in Figure 2.

2.1. Patient Information, Data Acquisition, and Segmentation. After excluding samples with incomplete data, a set of 30 patients was randomly selected from The Cancer Imaging Archive (TCIA, http://www.cancerimagingarchive.net/) publicly available database for a preliminary study. To obtain full imaging sets, 30 other GBM patients' data (age 50-68 years; 15 males, 15 females) were additionally chosen randomly from the TCIA database. Image pixels of the tumor regions were independently normalized on a scale from 0 to 1 (e.g., Figure 1(b) where the x-axis represents the normalized tumor pixels). The images were transformed into gray scale (using Matlab 2013 software) before further processing. Only FLAIR sequences were considered in this study. All the images were reconstructed to 512×512 matrices by segmenting the appropriate area of GBM by board certified radiologists, using the 3D slicer tool (Figure 3) [14]. Moreover, vE and vAT phenotypes were segmented manually slice-by-slice and organized in order to extract the statistical features. Statistical features of edema and active tumor parts were extracted then from raw FLAIR images.

2.2. Statistical Features Extraction. Features were extracted from the histogram shape, which is an area of the variable



FIGURE 1: Histogram of the GBM tumor. (a) Raw image of FLAIR sequence; (b) two Gaussian distributions represent vE and vAT and necrosis parts which are located inside the vAT with lower intensity values.

description based on the shape, and provided the frequency of values from different ranges of the variable. These features were applied previously in cervical cancer diagnosis using histogram based analyses of diffusion-weighted MR and its relation to histological features, subtype, and grade of cervical cancer [15, 16]. We quantified the two GBM phenotypes by nine statistical functions (Table 1).

All GBM patient data were plotted as histograms showing individual GBM data and their respective frequencies (Figure 1(b)). Features describing major statistical characteristics of these distributions were extracted according to Table 1 [15]. All features were extracted from histograms of GBM, according to

$$R_{\nu AT} = \begin{bmatrix} f_{1,1} & \cdots & f_{1,9} \\ \vdots & \ddots & \vdots \\ f_{m,1} & \cdots & f_{m,9} \end{bmatrix},$$
(1)

where *m* is the number of *v*AT samples, for each patient's one feature vector included nine features:

$$R_{\nu E} = \begin{bmatrix} f_{1,1} & \cdots & f_{1,9} \\ \vdots & \ddots & \vdots \\ f_{k,1} & \cdots & f_{k,9} \end{bmatrix},$$
 (2)



FIGURE 3: Example of phenotype segmentation. (a) Raw image of FLAIR sequence, (b) edema part vE, and (c) active tumor vAT.

where k is the number of vE, for each sample, and is similar in size to the nine features.

Note that the feature value represents the average of corresponding values of all slices in each patient.

One matrix vector R is organized according to

$$R = \{R_{vAT}, R_{vE}\}.$$
 (3)

For the GBM heterogeneity analysis, the aforementioned histogram features were extracted from the FLAIR MR images corresponding to the heterogeneity of vAT and vE. Therefore, the length of the resulting feature vector was nine. This statistical feature vector was taken as GBM heterogeneity based on vAT and vE, for the classification task at hand.

2.3. Statistical Analysis. Features were normalized using Z-scores which convert each of the feature vectors to have zero mean and unit variance. Moreover, an ANOVA test was used to assess the statistical significance between features and phenotypes [17]. This test was used to select the feature where p value < 0.01 was considered significant. Note that the total statistical features were found to be significant which are reported in Table 2.

2.4. Classifier Setting and Performance Metrics. Supervised technique such as the support vector machine (SVM) [18], naïve Bayes (NB) [19], and decision trees (DT) classifier [20] has become a popular learning algorithm for data mining applications, as employed to classify vAT from vE. A leave-one-out cross-validation was applied to obtain closely

Symbol	Features
f_1	Geometric mean, indicates the central tendency
f_2	Harmonic mean, calculates the average sample
f_3	Mean excluding outliers, measures the probability distribution
f_4	Mean (average)
f_5	Standard deviation (absolute deviation)
f_6	75th percentile, splits off the highest 25% of pixels from the lowest 75%
f_7	Quantile
f_8	Skewness, assesses the asymmetry of the distribution
f_9	Kurtosis, measures the degree of peakedness of a distribution
Features vector	$F = \{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8, f_9\}$

TABLE 2: Mean \pm standard deviation of *v*AT and *vE*.

Features	νAT	νE	<i>p</i> value
f_1	473.02 ± 345.65	461.15 ± 341.98	< 0.01
f_2	466.49 ± 344.81	453.57 ± 342.15	< 0.01
f_3	478.88 ± 347.18	468.25 ± 342.53	< 0.01
f_4	47.50 ± 31.33	53.69 ± 31.15	< 0.01
f_5	37.92 ± 25.98	44.19 ± 25.18	< 0.01
f_6	327.31 ± 294.16	321.74 ± 302.46	< 0.01
f_7	519.27 ± 367.27	516.73 ± 359.66	< 0.01
f_8	-0.20 ± 0.29	-0.09 ± 0.40	< 0.01
f_9	3.58 ± 0.85	2.93 ± 0.59	< 0.01

unbiased estimates of classification error rates. Additionally, receiver operating characteristics (ROC) curves and the corresponding areas under the ROC curve (AUC, with a cutoff value of 0.5), classifier accuracy, and confusion matrix were calculated to determine the performance of statistical feature for predicting the two GBM phenotypes.

Classifier accuracy measures the new sample correctly classified. It can be determined by the following expression:

Classifier accuracy

$$= \frac{vAT \text{ and } vE \text{ samples correctly classified}}{\text{Total number of samples}}$$
(4)
$$= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}},$$

where the true positive (TP) and the true negative (TN) are the number of *v*AT and *vE* samples correctly classified into positive and negative classes. The false positive (FP) and false negative (FN) are those samples which are incorrectly classified. Then, TP + FN are the total number of test samples of the considered class.

The results of the performance metrics reflect the value of this study in which the histogram (statistical) based features could be promising in discriminating between both types

TABLE 3: Metrics (%) of *v*AT and *vE* discrimination.

Features	DT		SVM		NB	
	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC
Full feature set (<i>F</i>)	68.33	96.05	68.33	80.22	53.33	77.66
Subset feature	75	92.5	58.33	73.88	58.33	76.44

TABLE 4: Confusion matrix based on selected features.

Features	D	Т	SV	М	N	В
	νAT	νE	νAT	νE	νAT	νE
30 vAT	20	10	15	15	13	17
30 vE	5	25	10	20	8	22

of GBM heterogeneity (vAT and vE). Due to this limited accuracy based on a full feature set, DT to find the optimal subset features was considered in order to improve the classifier accuracy. Simulation results were reported in Tables 3 and 4.

2.5. Features Selection Based on Decision Tree. Dominant features can be obtained using the decision structure "tree" model based on the general minimizing error. This model was proposed from various inducers, some comprising two conceptual phases "growing" and "pruning" (C4.5 [21] and CART [13]). The most important aspect of a decision tree induction strategy is the split criteria, which is the method of selecting an attribute that determines the distribution of training objects into subsets upon which subtrees are consequently built. The dominant features can be determined when subtrees are constructed. The choice for best attribute splitting was based on several techniques. This study used the Gini index $(I_{G(t)})$ as a more effective technique for splitting data and to detect the optimal subset features. $I_{G(t)}$ is an impurity-based criterion that measures the divergences between the probability distributions of the attribute's values. It is expressed as

$$I_{G(t)} = \sum_{i} p_i (1 - p_i),$$
 (5)

where p_i is the relative frequency of class *i* at node *t* and node *t* represents any node at which a given split is performed. p_i is determined by dividing the total number of observations in the class by the total number of observations.

DT was applied on the given dataset using a built-in *Matlab* function from the *decision trees for regression and classification* toolbox. Comparative results were reported in Tables 3 and 4.

3. Experimental Results

GBM phenotypes (vAT and vE) were segmented by board certified radiologists using a manual technique of 3D slicer tools. 30 vAT and 30 vE areas from raw MRI data of 30 patients were analyzed.

3.1. Feature Analysis. An ANOVA test showed that the nine statistical features were significant with p value < 0.01 (Table 2). Then, except for features f_4 , f_5 , and f_8 , features f_1 , f_2 , f_3 , f_6 , f_7 , and f_9 in vAT were significantly higher than the corresponding vE features. This is indicating that vAT is more pronounced statistically than vE.

3.2. Feature Selection. Feature selection was performed by decision tree model to determine the dominant statistical features, which would provide reliable discrimination between vAT and vE. Figure 4 shows the resulting decision tree. We observed that Kurtosis (f_9) and Skewness (f_8) play a dominant role as they appear towards the top of the tree structure. However, the other features have been identified as irrelevant attributes for the classification problem at hand since they do not appear in the tree.

3.3. Classification and Performance Comparison. A comparative study was done using the three-classifier model based on full feature set and subset feature. Table 3 shows 53.33– 68.33% range of accuracy classification using full features set and 58.33–75.00% for the subset feature with a highest value achieved using a decision tree classifier. Note that the SVM and NB classifier were not promising for discrimination between vAT and vE.

Moreover, AUC value shows a range of 77.66-96.05% for full feature set and 73.88-92.50% for subset feature with a highest value achieved using the decision tree classifier (Figure 5). This demonstrates the feasibility to discriminate between *v*AT and *vE* using the feature selection extracted from the FLAIR sequence.

Confusion matrix of the phenotypes discrimination showed that the 20 vAT and 25 vE samples are correctly classified from 30 phenotype samples based on DT classifier (Table 4). High rate of misclassification samples was in vAT samples where 10 of 30 samples were incorrectly classified as vE phenotype.

4. Discussion

In this study, we used statistical features extraction to assess vAT and vE phenotypes based on FLAIR sequence. The classifier accuracy of 75.00% was achieved using 30 patients and two features selected based on decision tree model. Among the nine features, Kurtosis (f_9) and Skewness (f_8) values might reflect the appropriate features which represented less correlation with other features. In this context, Figure 6 shows the heat map of the correlation coefficients between the nine features. We observed that the lowest correlation coefficients were achieved by two features Kurtosis (f_9) and Skewness (f_8) . Note that the higher correlation coefficients represent the common characters between vE and vAT, as shown in the statistical features $(f_1, f_2, f_3, f_4, f_5, f_6, and f_7)$. These two selected features can be associated with the phenotype heterogeneity.

In this context, multiple studies have suggested that increasing heterogeneity is associated with cancer [22]. The results demonstrated that building features and biological



FIGURE 4: Decision tree grown using 9 statistical features extracted from 30 vAT and 30 vE parts.



FIGURE 5: Receiver operating characteristic curves for distinguishing between *v*AT and *vE*. FFS denotes full feature set, and FS is the feature selection.

significance are promising for noninvasive detection of GBM heterogeneity based on *v*AT and *vE*.

Moreover, recent efforts concluded that future research will be most productive by focusing on genetics, clinical data, and imaging features [23]. Thus, characterizing the molecular properties of GBM and making them publicly available are goals of the Cancer Genome Atlas (TCGA) [24].

Traditionally, the feature extraction applied to medical imaging was limited to the whole tumor and normal brain [23]; however, current advances in medical image processing like the findings presented in this study allow for highthroughput extraction of characteristic imaging features to



FIGURE 6: Heat map with correlation coefficients between statistical features.

measure complex and very subtle differences across patient MRIs. Thus, these findings provide strong evidence that feature extraction can identify and discriminate between GBM heterogeneity types.

Previous study was done to apply texture analysis for assessment of the traumatic brain injury [25] and also to discriminate between the GBM phenotypes; however, two MRI sequences were used, features of necrosis and active tumor parts from T1-WI images and edema parts from FLAIR sequence. The texture feature extracted from GLCM was considered, and the simulation results for 13 patients show the highest accuracy of 67% [26].

Obviously, neuroradiologists are becoming increasingly important players for early diagnosis of GBM. Our vision is to integrate engineering based methods as described in daily practice to enhance radiologists' performance beyond their routine "vision." Particularly, in utterly devastating disease like GBM, improvements in any medical specialty involved are of the utmost essence. Note that numerous factors may have led to varying results between this study and previously published studies potentially due to the following reasons.

In this work, the whole GBM tumor was assessed. Only FLAIR sequence was considered and only two phenotypes were addressed based on the data distribution (Figure 1). It would have been preferable to include more patients to strengthen the GBM heterogeneity analysis. However, the number of patients included in this study can provide preliminary information about GBM heterogeneity.

5. Conclusions

This paper analyzed and implemented GBM vAT and vE discrimination based on the statistical features extracted from MRI raw data. For the analysis of GBM heterogeneity, feature extraction was more effective as it could characterize each phenotype by a specific set of features to robustly

identify them. By automatic recognition, this identification subsequently provided a more accurate assessment of the patient prognosis and underlying genomic composition. Improved classifier accuracy was achieved using the decision tree model. Feature extraction, selection, learning, and classification were applied on 30 vAT and 30 vE phenotypes.

The experimental results were confirmed by higher accuracy classifier of appropriate features based on two features Kurtosis and Skewness. The drop in average correct classification rate resulted from difficulty in classifying based on histogram and subset feature. Histogram feature extracted from GBM phenotypes yields a promising technique for differentiating *v*AT from *vE* indicating the oncological level of aggressiveness of a tumor. Extending this work by increasing the number of patients would enhance the accuracy of GBM heterogeneity prediction in the future.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by the University of Lorraine.

References

- R. Stupp, M. E. Hegi, M. J. van den Bent et al., "Changing paradigms—an update on the multidisciplinary management of malignant glioma," *The Oncologist*, vol. 11, no. 2, pp. 165–180, 2006.
- [2] B. J. McCarthy, C. Kruchko, and T. A. Dolecek, "The impact of the Benign Brain Tumor Cancer Registries Amendment Act (Public Law 107–260) on non-malignant brain and central nervous system tumor incidence trends," *Journal of Registry Management*, vol. 40, no. 1, pp. 32–35, 2013.
- [3] E. D. Angelini, O. Clatz, E. Mandonnet, E. Konukoglu, L. Capelle, and H. Duffau, "Glioma dynamics and computational models: a review of segmentation, registration, and in silico growth algorithms and their clinical applications," *Current Medical Imaging Reviews*, vol. 3, no. 4, pp. 262–276, 2007.
- [4] W. B. Pope, J. R. Young, and B. M. Ellingson, "Advances in MRI assessment of gliomas and response to anti-VEGF therapy," *Current Neurology and Neuroscience Reports*, vol. 11, no. 3, pp. 336–344, 2011.
- [5] A. Chaddad, C. Tanougast, A. Golato, and A. Dandache, "Carcinoma cell identification via optical microscopy and shape feature analysis," *Journal of Biomedical Science and Engineering*, vol. 6, no. 11, pp. 1029–1033, 2013.
- [6] A. Chaddad, "Automated feature extraction in brain tumor by magnetic resonance imaging using gaussian mixture models," *International Journal of Biomedical Imaging*, vol. 2015, Article ID 868031, 11 pages, 2015.
- [7] A. Chaddad, C. Tanougast, A. Dandache, and A. Bouridane, "Extracted haralick's texture features and morphological parameters from segmented multispectrale texture bio-images for classification of colon cancer cells," WSEAS Transactions on Biology and Biomedicine, vol. 8, no. 2, pp. 39–50, 2011.

- [8] W. P. Kegelmeyer Jr., "Computer detection of stellate lesions in mammograms," in *Biomedical Image Processing and Three-Dimensional Microscopy*, vol. 1660 of *Proceedings of SPIE*, pp. 446–454, San Jose, Calif, USA, February 1992.
- [9] B. L. Wen, M. A. Brewer, O. Nadiarnykh et al., "Texture analysis applied to second harmonic generation image data for ovarian cancer classification," *Journal of Biomedical Optics*, vol. 19, no. 9, Article ID 096007, 2014.
- [10] M. F. Ahmad Fauzi, H. N. Gokozan, B. Elder, V. K. Puduvalli, J. J. Otero, and M. N. Gurcan, "Classification of glioblastoma and metastasis for neuropathology intraoperative diagnosis: a multi-resolution textural approach to model the background," in *Medical Imaging 2014: Digital Pathology*, vol. 9041 of *Proceedings of SPIE*, p. 9, March 2014.
- [11] R. M. Palenichka, M. B. Zaremba, and R. Missaoui, "Multiscale model-based feature extraction in structural texture images," *Journal of Electronic Imaging*, vol. 15, no. 2, Article ID 023013, 2006.
- [12] A. Chaddad and R. R. Colen, "Statistical feature selection for enhanced detection of brain tumor," in *Applications of Digital Image Processing XXXVII*, vol. 9217 of *Proceedings of SPIE*, p. 8, 2014.
- [13] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*, Chapman and Hall/CRC, New York, NY, USA, 1st edition, 1984.
- [14] 3D Slicer, 2014, http://www.slicer.org/.
- [15] K. Downey, S. F. Riches, V. A. Morgan et al., "Relationship between imaging biomarkers of stage I cervical cancer and poor-prognosis histologic features: quantitative histogram analysis of diffusion-weighted MR images," *American Journal of Roentgenology*, vol. 200, no. 2, pp. 314–320, 2013.
- [16] A. B. Rosenkrantz, "Histogram-based apparent diffusion coefficient analysis: an emerging tool for cervical cancer characterization?" *American Journal of Roentgenology*, vol. 200, no. 2, pp. 311–313, 2013.
- [17] A. Cuevas, M. Febrero, and R. Fraiman, "An anova test for functional data," *Computational Statistics and Data Analysis*, vol. 47, no. 1, pp. 111–122, 2004.
- [18] M. A. Hearst, S. T. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and Their Applications*, vol. 13, no. 4, pp. 18–28, 1998.
- [19] C. C. Aggarwal, Data Classification: Algorithms and Applications, CRC Press, New York, NY, USA, 2014.
- [20] L. Rokach, *Data Mining with Decision Trees: Theory and Applications*, World Scientific, River Edge, NJ, USA, 2007.
- [21] J. R. Quinlan, *C4.5: Programs for Machine Learning*, Morgan Kaufmann, San Francisco, Calif, USA, 1993.
- [22] R. A. Burrell, N. McGranahan, J. Bartek, and C. Swanton, "The causes and consequences of genetic heterogeneity in cancer evolution," *Nature*, vol. 501, no. 7467, pp. 338–345, 2013.
- [23] S. Herlidou-Même, J. M. Constans, B. Carsin et al., "MRI texture analysis on texture test objects, normal brain and intracranial tumors," *Magnetic Resonance Imaging*, vol. 21, no. 9, pp. 989– 993, 2003.
- [24] Cancer Genome Atlas Research Network, "Comprehensive genomic characterization defines human glioblastoma genes and core pathways," *Nature*, vol. 455, no. 7216, pp. 1061–1068, 2008.
- [25] K. K. Holli, L. Harrison, P. Dastidar et al., "Texture analysis of MR images of patients with Mild Traumatic Brain Injury," *BMC Medical Imaging*, vol. 10, no. 1, article 8, 2010.

[26] A. Chaddad, P. O. Zinn, and R. R. Colen, "Quantitative texture analysis for Glioblastoma phenotypes discrimination," in *Proceedings of the International Conference on Control, Decision and Information Technologies (CoDIT '14)*, pp. 605–608, Metz, France, November 2014.