Predicting the Molecular Subtypes in Gliomas Using T2-Weighted MRI

Jayant Jagtap, Jitender Saini, Vani Santosh and Madhura Ingalhalikar

Abstract Classifying gliomas noninvasively into their molecular subsets is a crucial neuro-scientific problem. Prognosis of the isocitrate dehydrogenase (IDH) mutation in gliomas is important for planning targeted therauptic intervention and tailored treatment for individual patients. This work proposes a novel technique based on texture analysis of T2-weighted magnetic resonance imaging (MRI) scans of grade 2 and grade 3 gliomas to differentiate between IDH1 mutant 1p/19q positive and IDH1 mutant 1p/19q negative categories. The textural features used in the proposed method are local binary patterns histogram (LBPH), Shannon entropy, histogram, skewness, kurtosis, and intensity grading. We discriminate the tumors into their molecular subtypes using standard artificial neural networks (ANNs). LBPH attributes demonstrated maximum discrimination between the two groups followed by Shannon entropy. In summary, the technique proposed facilitates an early biomarker to detect the IDH subtype noninvasively and can be employed as an automated tool in clinics to aid diagnosis.

Keywords Isocitrate dehydrogenase (IDH) · Gliomas · Magnetic resonance imaging (MRI) · Local binary patterns histogram (LBPH) · Artificial neural networks (ANN)

J. Jagtap (✉) · M. Ingalhalikar
Symbiosis Institute of Technology, Symbiosis International University,
Pune 412115, India
e-mail: jayant.jagtap@sitpune.edu.in

M. Ingalhalikar
e-mail: madhura.ingalhalikar@sitpune.edu.in

J. Saini · V. Santosh
National Institute for Mental Health and Neurosciences, Bengaluru 560029, India
e-mail: jitender.s@nimhans.kar.nic.in

V. Santosh
e-mail: vani@nimhans.kar.nic.in

A. J. Kulkarni et al. (eds.), *Proceedings of the 2nd International Conference on Data Engineering and Communication Technology*, Advances in Intelligent Systems and Computing 828, https://doi.org/10.1007/978-981-13-1610-4_7

1 Introduction

Gliomas are the most common type of tumor of the central nervous system that are usually life threatening and require early diagnosis for better mortality [1, 2]. Traditionally tumors are classified into grade I to IV subtypes that are based on the microscopic features provided via biopsy procedure. With the advent of computer-aided diagnosis (CAD), the tumors can be marked using magnetic resonance imaging (MRI). The location and size can be accurately mapped; however, the inhomogeneity in intensity and high heterogeneity in each subtype makes it difficult to discriminate the tumors. However, recent genetic molecular advances have contributed to a better understanding of tumor pathophysiology as well as disease stratification. Recent studies have identified novel mutations in the isocitrate dehydrogenase (IDH)1 gene from a genome-wide mutational analysis of glioblastomas [3]. Further, it has been found that IDH1 mutations are present in 50–90% of cases of grade II and III gliomas and are associated with longer survival in glioma patients [4]. However, to identify the IDH1 status, a genome-wide study is required.

The aim of this research is to present an automated noninvasive CAD approach that can categorize brain tumors MRI scans into their molecular subtypes namely oligodendroglioma IDH1 mutant 1p/19q positive and anaplastic astrocytoma IDH1 mutant 1p/19q negative. The proposed approach can be employed as an early molecular biomarker and can therefore aid in the diagnosis and treatment planning of the patient.

Recent work on characterizing IDH1 subtypes includes textural analysis of T2-weighted images by using Shannon entropy and edge maps using Prewitt filtering [5]. The study demonstrated IDH1 wild-type gliomas to have statistically lower Shannon entropy than IDH1-mutated gliomas; however, no significant difference was observed on the edge strength maps. Another recent study employed multimodal approach by using T2-weighted and diffusion MRI images [6]. Texture features that include entropy, skewness, and kurtosis were computed. However, other than entropy none of the features could segregate the molecular subtypes. This work is highly innovative, as we propose to employ multiple texture features that include the local binary patterns histogram (LBPH), Shannon entropy, histogram, skewness, kurtosis, and intensity grading, which provide superior discrimination between the 1p/19q positive and 1p/19q negative molecular subgroups. The classification is achieved by using a simple two-layer feedforward backpropagation artificial neural network (ANN). Despite the inhomogeneity in intensity levels and variability in size and structure, we obtain a classifier with high accuracy, establishing the applicability for early diagnosis of the gliomas.

2 Proposed Method for Glioma Classification

2.1 Overview

Our method of creating texture-based glioma classifiers includes the following steps: (1) manual segmentation of gliomas on T2-weighted images, (2) feature extraction, (3) classifier training and cross-validation. Figure 1 displays a schematic block diagram of the method. In the next few sections, we shall first describe the features involved and the classification technique, followed by the experiment on the T2-weighted images.

2.2 Feature Extraction

The textural features used in the proposed method are local binary patterns histogram (LBPH), Shannon entropy, histogram, skewness, kurtosis, and intensity grading.

1. **Local Binary Patterns**
   Local binary patterns (LBPs) proposed by Ojala et al. [7] are further grouped into 10 codes from 0 to 9 [8]. The histogram of these 10 LBP codes, called as local binary pattern histogram (LBPH), is used as textural features in the proposed method.
2. **Shannon Entropy**
   Shannon entropy is the amount of information carried by an image. The Shannon entropy E is calculated by using Eq. 1 as follows.

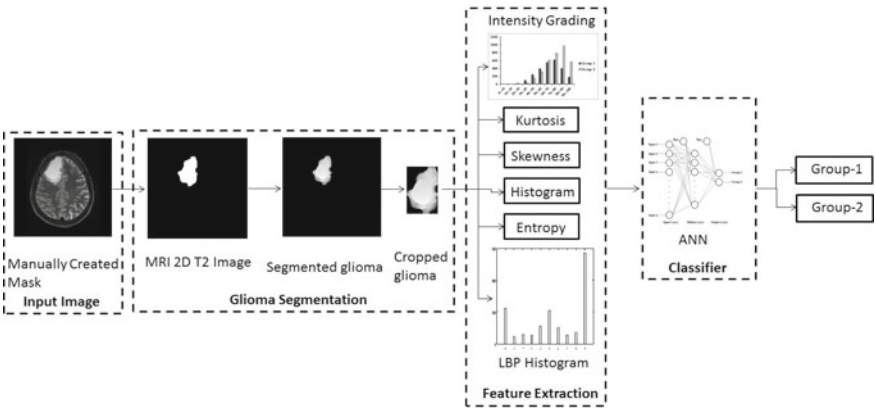

Fig. 1 Work flow diagram of proposed method for glioma classification

$$E = -\sum_{g=0}^{255} f_g log_2(f_g) \tag{1}$$

where f_g denotes the frequency of occurrence of gth the gray level.

3. **Histogram, Skewness, Kurtosis**
   Histogram gives the frequency of occurrence of a particular gray level in an image. Skewness and kurtosis are the third and fourth moment of data distribution, respectively.
4. **Intensity Grading**
   Intensity grading gives the count of number of pixels lies within the specified range of intensities in an image. Intensity grading is also calculated from histogram of an image.

2.3 Classifier

The artificial neural network (ANN) classifier is used to classify the gliomas into one of the two groups. ANN is designed by using multilayer feedforward backpropagation neural networks [9]. It consists of an input layer, a hidden layer, and an output layer. Input layer consists of number of neurons equal to the length of feature vector of an input image. Output layer consists of two neurons excluding bias neuron equal to number of output groups. For the hidden layer, we select number of neurons equal to $\frac{1}{3}$rd of the neurons used in the input layer.

2.4 Algorithm

The steps involved in training and testing of the proposed method are as follows.

1. Let $I(x, y)$ be an input grayscale image, where x and y are the pixel intensities in the range [0, 255].
2. Resize $I(x, y)$ to 256×256.
3. Manually mark glioma in $I(x, y)$.
4. Let $I_C(x, y)$ be the cropped glioma image.
5. Perform texture analysis for feature extraction using LBPH and entropy.
   a. Calculate LBPH of I_C and store in a vector F_{LBPH}, as given in Eq. 2.

$$F_{LBPH} = [LBPH(I_C)]_{10\times 1} \tag{2}$$

   b. Calculate the entropy of I_C and store in a vector F_{ENT}, as given in Eq. 3.

$$F_{ENT} = [Entropy(I_C)]_{1\times 1} \tag{3}$$

c. Calculate the histogram of I_C and store in a vector F_{HIST}, as given in Eq. 4.

$$F_{HIST} = [Histogram(I_C)]_{255\times 1} \tag{4}$$

d. Calculate skewness of I_C using F_{HIST} and store in a vector F_{SKEW}, as given in Eq. 5.

$$F_{SKEW} = [Skewness(I_C)]_{1\times 1} \tag{5}$$

e. Calculate kurtosis of I_C using F_{HIST} and store in a vector F_{KURT}, as given in Eq. 6.

$$F_{KURT} = [Kurtosis(I_C)]_{1\times 1} \tag{6}$$

f. Calculate intensity grading of I_C using F_{HIST} and store in a vector F_{IG}, as given in Eq. 7

$$F_{IG} = [IntensityGrading(I_C)]_{10\times 1} \tag{7}$$

g. Concatenate all the textural feature vectors calculated in steps a to f in a vector F_T, called textural feature vector, as given in Eq. 8.

$$F_T = [F_{LBPH}, F_{ENT}, F_{HIST}, F_{SKEW}, F_{KURT}, F_{IG}]_{278\times 1} \tag{8}$$

h. The normalized feature vector F_{T_n} is calculated by normalizing vector F_T using min-max normalization technique as given in Eq. 9.

$$F_n = \frac{F_{a_n} - min(F_a)}{max(F_a) - min(F_a)} \tag{9}$$

where $F_a = (F_{a_1}, ..., F_{a_n})$, n is the length of feature vector, and F_n is n^{th} normalized data.

6. Apply the normalized feature vector F_{T_n} to ANN classifier, with corresponding group labels in training phase, and without corresponding group labels in testing phase, to classify the input images into one of the two groups, namely Group-1(1p/19q positive) and Group-2(1p/19q negative).

3 Experimental Results

3.1 Dataset

The glioma dataset was acquired in part of the effort at National Institute of Mental Health And Neuro-Sciences (NIMHANS), Bangalore. All the patients were clinically assessed by an experienced oncologist. The subjects, 15 with IDH1 1p/19q positive and 10 with IDH1 1p/19q negative, were recruited. These subjects were in

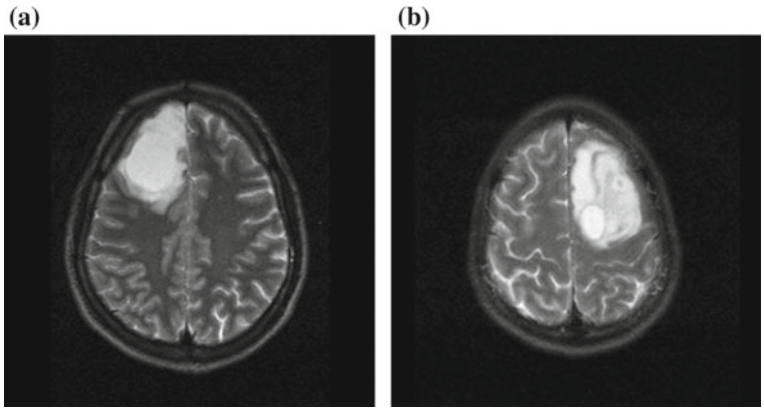

Fig. 2 Sample images from dataset **a** Group-1, **b** Group-2

the age range of 17 years to 58 years with an average age of 38 years. All imaging was performed on the same site, using the Siemens 1.5 T with a 12 channel head coil. The T2-weighted images were acquired with the following parameters: TR/TE = 4500/90 ms, slice thickness = 5 mm, and voxel-size = 1 × 1 × 5 mm. Since these images acquired at a low resolution in the sagittal and coronal planes, we used only the 2D axial slices in this work.

We employed 18 axial slices from 15 subjects in the 1p/19q positive group and 15 axial slices from the 10 subjects in 1p/19q negative group. These slices were chosen at the maximum cross-sectional area of the tumor, with each patient contributing not more than 3 slices. Figure 2 demonstrates the sample slices for the 1p/19q positive and 1p/19q negative groups. The figure is shown to illustrate the complexity of the classification problem.

3.2 Feature Extraction

By using the steps mentioned in Sect. 2, first, the LBPH of glioma-segmented image is calculated and stored in a vector F_{LBPH} of size 10 × 1. Then, entropy of the glioma segmented image is calculated and stored in a vector F_{ENT} of size 1 × 1. Next, histogram of the glioma-segmented image is calculated and stored in a vector F_{HIST} of size 255 × 1. Later on, skewness and kurtosis of the glioma-segmented image are calculated and stored in vector F_{SKEW} of size 1 × 1 and F_{KURT} of size 1 × 1, respectively. But last, intensity grading of glioma-segmented image is calculated and stored in a vector F_{IG} of size 10 × 1. Finally, all the features are concatenated and stored in a vector F_T of size 278 × 1.

3.3 Classification

The feature vector F_T is normalized by using min-max normalization technique, as discussed in section 2, to get the values of the features in the range [0 1]. The normalized feature vector F_{T_n} is applied to ANN classifier with corresponding group labels in training phase and without corresponding group labels in testing phase to classify the input images into one of the two groups. The ANN is designed by using two-layer feedforward backpropagation algorithm. The ANN consists of an input layer, a hidden layer, and an output layer. The input layer consists of 278 neurons, hidden layer consists of 93 neurons, and the output layer consists of 2 neurons. The values for rest of the parameters of ANN such as number of epochs, training function, error function are fixed experimentally to get maximum classification accuracy.

3.4 Statistical Analysis

Statistical analysis is carried out to measure the performance of the proposed method. The statistical analysis using *t*-test is carried out for comparing two groups. A Receiver Operating Characteristic (ROC) curve is plotted in turn which shows sensitivity and specificity for estimated groups with respect to predefined groups, as shown in Fig. 3. The Area Under the Curve (AUC) of 0.8993 is obtained from ROC. Sensitivity and specificity of 83.3% and 93.3%, respectively, are achieved by using proposed method to classify the gliomas into two groups. A *p*-value of <0.05 is considered statistically significant. A *p*-value of 0.0107 is observed between the feature vectors of two groups. LBPH, entropy, and intensity grading show a significantly lower *p*-value (<0.05), whereas histogram, skewness, and kurtosis demonstrate insignificant results. Experimentally, it has been observed that LBPH is the most efficient textu-

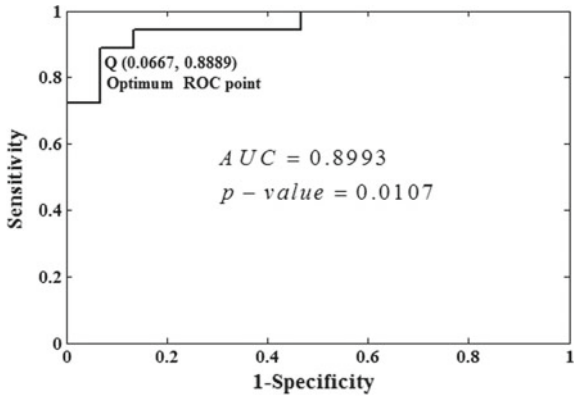

Fig. 3 ROC for glioma classification into two groups using proposed method

ral feature to discriminate between Group-1 and Group-2 followed by entropy and intensity grading.

The classification accuracy of proposed method is calculated by using tenfold cross-validation. Overall classification accuracy of 87.9% is achieved by using proposed method to classify the glioma images into one of the two groups.

4 Conclusion and Future Scope

We have presented a classification methodology based on meaningful texture features for identification of molecular classes in gliomas. Superior experimental results, with high classification accuracy indicate that our technique, can be successfully employed in computed aided diagnostic systems. Future work includes extending and testing the classifier on a bigger dataset and not restricting the analysis to T2-weighted images, but using multimodal features from other MRI contrasts such as T1-weighted images, T1-contrast enhanced images, fluid-attenuated inversion recovery (FLAIR), and diffusion MRI.

In summary, the paper presents a proof of concept for noninvasively discriminating the gliomas into molecular subtypes and has the potential to be used as a diagnostic biomarker for IDH1.

Ethical Approval

All the subjects had provided written informed consent for their participation in this study and the Institute Ethics Committee of NIMHANS had approved this study.

References

1. Munshi A (2016) Central nervous system tumors: spotlight on India. South Asian J Cancer 5(3):146–147
2. Thakkar J, Dolecek T, Horbinski C et al (2014) Epidemiologic and molecular prognostic review of glioblastoma. Cancer Epidemiol Biomarkers Prev 23(10):1985–1996
3. Cohen A, Holmen S, Colman H (2013) IDH1 and IDH2 mutations in gliomas. Curr Neurol Neurosc Rep 13(5):345
4. Qi S, Yu L, Li H, et al (2014) Isocitrate dehydrogenase mutation is associated with tumor location and magnetic resonance imaging characteristics in astrocytic neoplasms. Oncol Lett 7(6):1895–1902
5. Kinoshita M, Sakai M, Arita H et al (2016) Introduction of high throughput magnetic resonance T2-weighted image texture analysis for WHO grade 2 and 3 gliomas. PLoS ONE 11(10):e0164268
6. Ryu Y, Choi S, Park S et al (2014) Glioma: application of whole-tumor texture analysis of diffusion-weighted imaging for the evaluation of tumor heterogeneity. PLoS ONE 9(9):e108335
7. Ojala T, Pietikainen M, Maenpaa T (2002) Multiresolution gray-scale and rotation Invariant texture classification with local binary patterns. IEEE Trans Pattern Anal Mach Intell 24(7):971–987

8. Maenpaa T, Pietikainen M (2004) Texture analysis with local binary patterns. Handbook of pattern recognition and computer vision. World Scientific, Singapore
9. Zurada JM (1992) Introduction to artificial neural systems. West Publishing Company, St. Paul, US