A New Post-processing Method to Detect Brain Tumor Using Rough-Fuzzy Clustering

Shaswati Roy and Pradipta Maji(✉)

Biomedical Imaging and Bioinformatics Lab, Machine Intelligence Unit,
Indian Statistical Institute, Kolkata, India
{shaswatiroy_t,pmaji}@isical.ac.in

Abstract. Automatic and accurate brain tumor segmentation from MR images is one of the important problems in cancer research. However, the lack of shape prior and weak contrast at boundary make unsupervised brain tumor segmentation more challenging. In this background, a new brain tumor segmentation method is being developed, integrating judiciously the merits of multiresolution image analysis technique and rough-fuzzy clustering. One of the major issues of the clustering based segmentation method is how to extract brain tumor accurately, since tumors may not have clearly defined intensity or textural boundaries. In this regard, this paper presents a new post-processing method for clustering based brain tumor detection. It combines the merits of mathematical morphology and the concept of rough set based region growing approach to refine the result obtained after clustering, thereby ensuring the accurateness of brain tumor segmentation application. The performance of the proposed approach, along with a comparison with related methods, is demonstrated on a set of synthetic and real brain MR images.

1 Introduction

Automatic segmentation of healthy and pathologic brain tissues from MRI plays an important role in brain tumor detection application. Early and accurate brain tumor segmentation from MR images is a difficult task in many cancer research applications. Brain tumors may be of any size, may have a variety of shapes, may appear at any location in brain, and may appear in different image intensities. Some tumors also deform other structures and appear together with edema that changes intensity properties of the nearby region.

The challenges associated with automatic brain tumor segmentation have given rise to many different approaches [2,7,8,10–12]. Classification based tumor detection algorithms are widely used in brain tumor detection applications. These methods are constrained to the supervised [7,9] or unsupervised [1]. Menze et al. [7] combined a healthy brain atlas with a tumorous brain atlas to segment tumors using a generative probabilistic model and spatial regularization. Bauer et al. [9] combined support vector machine using multispectral intensities and textures with subsequent hierarchical regularization based on

M. Kryszkiewicz et al. (Eds.): PReMI 2015, LNCS 9124, pp. 407–417, 2015.
DOI: 10.1007/978-3-319-19941-2_39

conditional random fields. This method requires four modalities, namely, T1-weighted, T1-weighted with contrast agent, T2-weighted, and FLAIR to classify the tumor region. Fletcher-Heath et al. [1] combined fuzzy clustering and integrated domain knowledge to improve the tumor segmentation applied on T1-, T2-, and PD-weighted images. A block diagram illustrating the main steps during the clustering based brain tumor segmentation pipeline is shown in Fig. 1. The pre-processing step may include denoising, skull stripping, correction of intensity inhomogeneity, or registration for multi-modal input images. Feature vectors are then generated for clustering algorithm to extract the tumor class.

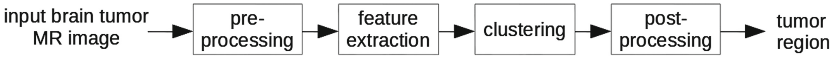


Fig. 1. Block diagram of the clustering based brain tumor detection method

In this regard, a texture-based brain MR image segmentation method is presented in [6] that can be used for detecting brain tumor from MR images. It judiciously integrates the merits of multiresolution image analysis and rough-fuzzy clustering. The multiresolution wavelet analysis is used to extract scale-space feature vector for each pixel of the given brain MR image. Since the boundary between brain and skull is relatively strong on T1 scan, a skull stripping algorithm is used to extract the brain tissues and remove non-cerebral tissues. However, the use of wavelet decomposition may give rise to some irrelevant and insignificant features. Hence, an unsupervised feature selection method is used to reduce the dimensionality of feature space by maximizing both relevance and significance of the selected features. Finally, the robust rough-fuzzy c-means algorithm [5] is used for segmentation of brain MR image. While the integration of both membership functions of fuzzy sets enables efficient handling of overlapping classes in noisy environment, the concept of lower and upper approximations of rough sets deals with uncertainty, vagueness, and incompleteness in class definition. In effect, it groups similar textured tissue classes contained in the image.

However, the lack of intensity or shape priors and weak contrast at boundary make unsupervised brain tumor segmentation more challenging. Once the tumor class is extracted by clustering algorithm, the result can be further improved by a sequence of post-processing steps as shown in Fig. 1. Hsieh et al. [2] proposed a simple region-growing and knowledge-based post-processing step after the fuzzy classification applied on non-contrasted T1- and T2-weighted MR images, followed by a morphology operator. Iscan et al. [3] developed a filtering and region growing based post-processing approach on segmented tumorous tissue.

In this background, this paper presents a new post-processing method. It integrates the merits of morphological operations and the notion of rough set theory in formulation of region growing method. Since tumors may not have clearly defined intensity or textural boundaries, there may be some ambiguity at boundary region of a tumor class. This uncertainty is handled by incorporating the rough set theory into region growing approach. Hence, this post-processing step plays an important role in order to ensure the correctness of the diagnosis in brain tumor segmentation applications. The performance of the proposed approach, along with a comparison with related methods, is demonstrated on a set of synthetic and real brain MR images both qualitatively and quantitatively.

2 Proposed Post-Processing Method for Tumor Detection

This section briefly presents rough-fuzzy clustering and wavelet based brain MR image segmentation method reported in [6]. It consists of five steps as follows:

1. Generation of mask from brain MR image for identification of brain region;
2. Dyadic wavelet analysis of MR image using Daubechies 6-tap filter;
3. Generation of feature vectors for brain region using the mask;
4. Unsupervised feature selection to select relevant and significant features for clustering; and
5. Rough-fuzzy clustering to generate segmented image.

However, the above algorithm results in oversegmentation or undersegmentation of brain tumor due to the ambiguities at tumor boundaries. Hence, a new post-processing method, based on morphology and rough set based region growing method is used to handle efficiently these ambiguities. Due to partial volume effect, the brain tumor segmentation based on clustering algorithm may produce many residual areas around the tumor region. Again, since the tumor is not homogeneous, clustering based segmentation may produce small holes within the tumor class. Hence, the morphological operations, namely, closing and opening, in the order, followed by finding largest connected component, are applied to eliminate these residual areas, as well as to eliminate any discontinuity within tumor mass. The closing and opening morphology use square shaped structuring element of different dimensions. Two-pass connected component labelling is embedded for finding the largest connected component. Then, a rough set based region growing method is proposed to efficiently handle the ambiguities at anatomical tumor boundaries.

The theory of rough sets begins with the notion of an approximation space, which is a pair $< U, R >$, where U be a non-empty set, the universe of discourse, and R an equivalence relation on U. Given an arbitrary set $X \in 2^U$, in general, it may not be possible to describe X precisely in $< U, R >$. One may characterize X by a pair of lower and upper approximations. The lower approximation $\underline{R}(X)$ is the union of all the elementary sets which are subsets of X, and the upper approximation $\overline{R}(X)$ is the union of all the elementary sets which have a non-empty intersection with X. The interval $< \underline{R}(X), \overline{R}(X) >$ is the representation of X in the approximation space $< U, R >$ or simply called the rough set of X. The accuracy of X, denoted by $\alpha_R(X)$, is the ratio of the number of objects in its lower approximation to that in its upper approximation; namely

$$\alpha_R(X) = |\underline{R}(X)||\overline{R}(X)|^{-1} \tag{1}$$

Note that the higher the accuracy of a subset, the better is its approximation.

The proposed region growing technique is based on the concept of rough set theory. Let T be the tumor class. Application of various morphological operations produces the region that possibly belongs to T, that is, upper approximation of the tumor class, $\overline{A}(T)$. The lack of shape prior and weak contrast at anatomical boundary induce uncertainty of belongingness of boundary region pixels. Hence, it puts some non-tumorous cells into $\overline{A}(T)$ because they are characterized as indiscernible with other actual tumorous pixels, using the available information.

The accuracy of approximation of T as in (1), in fact, provides a measure of how closely the rough set is approximating the target set with available knowledge. Hence, the non-tumorous tissues in $\overline{A}(T)$ are discarded in this methodology in order to increase the accuracy of approximation. In this regard, a region growing approach based on the gradient information of input image is proposed, which concentrates on elimination of non-tumorous healthy brain tissues from pathologic region appropriately, thereby maximizing the rough set accuracy.

The center of gravity (CoG) is computed within $\overline{A}(T)$. The CoG of $\overline{A}(T)$ is assumed to certainly belong to the tumor region, and hence the CoG is assigned initially in the lower approximation of the tumor class $\underline{A}(T)$. Then, a simple region growing approach is followed with CoG as the seed point. An edge map of the input brain MR image is constructed within the region of $\overline{A}(T)$. In this approach, Sobel's gradient operator is used. It starts growing initially from the CoG to spread over the edge map of $\overline{A}(T)$. The region of interest $\underline{A}(T)$ expands until a certain criterion is met which is based on the mean gradient magnitude of input image within the region of $\overline{A}(T)$.

Let $\delta(x)$ be gradient magnitude of pixel x. The threshold Δ is computed as:

$$\Delta = \frac{1}{\left|\overline{A}(T)\right|} \sum_{x \in \overline{A}(T)} \delta(x), \tag{2}$$

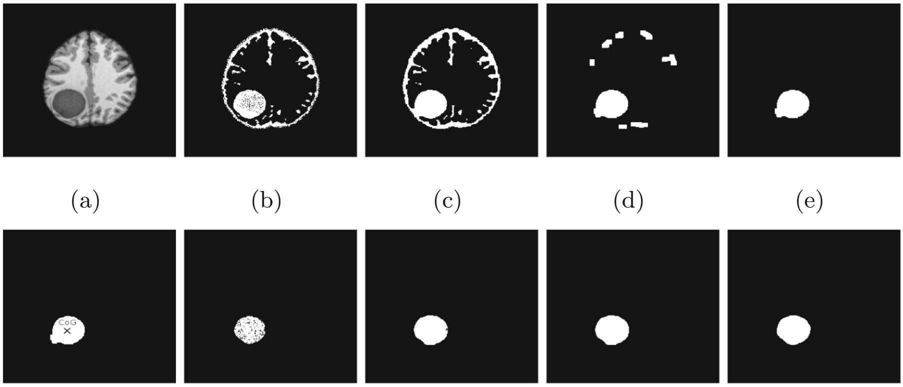
Based on the value of Δ, the unassigned pixel is included in $\underline{A}(T)$ using the pixel assignment rule for lower approximation described as below:

$$\underline{A}(T) \leftarrow \underline{A}(T) \cup \left\{x \in \overline{A}(T) | \delta(x) \leq \Delta \wedge (\exists y \in \underline{A}(T) \text{ s.t. } x \in \mathcal{N}(y))\right\}, \tag{3}$$

where $\mathcal{N}(p)$ represents the neighbors of the pixel p using eight-connectivity. This region growing method is repeated until no more changes occurred. After building $\underline{A}(T)$, the boundary of $\underline{A}(T)$ is refined using the lower approximation refinement rule as presented below:

$$\underline{A}(T) \leftarrow \underline{A}(T) \cup \left\{x \in \overline{A}(T) | \delta(x) \geq \Delta \wedge (\exists y \in \underline{A}(T) \text{ s.t. } x \in \mathcal{N}(y))\right\}. \tag{4}$$

This assignment rule includes the pixels that represent the boundary edge region of tumor. Figure 2 shows an example of building the rough approximation $\overline{A}(T)$ and 2D region growing approach applied on CoG of $\overline{A}(T)$ to approximate the corresponding tumor class. As it can be seen in the tumor result of Fig. 2g,



(a) (b) (c) (d) (e)

(f) (g) (h) (i) (j)

Fig. 2. An example of detection of tumor region T: (a) input, (b) after clustering, (c) closing, (d) opening, (e) largest connected component, i.e., $\overline{A}(T)$, (f) CoG marking, (g) $\underline{A}(T)$, (h) $\underline{A}(T)$ refinement, (i) final closing, and (j) ground truth

there are several holes inside the brain region caused by the higher gradient magnitude values for some pixels within tumor region than the threshold Δ computed, due to its noise factor. Also, the gradient magnitude on tumor edge boundaries may reach above the threshold value Δ. These are refined using the lower approximation refinement rule of (4). Finally, the morphological closing operation is used to smooth the surface of the tumor class. The main steps of the proposed morphology and rough set based post-processing method proceed as follows:

1. Apply closing morphology with a window of dimension 5×5 on the tumor class extracted by clustering algorithm.
2. Apply opening morphology with a window of size 7×7 in order to separate the tumor region from healthy brain tissues.
3. Find the largest connected component using connected component labelling and consider it as rough approximation of tumor $\overline{A}(T)$.
4. Construct an edge map from input brain MR image for region of $\overline{A}(T)$.
5. Compute the threshold value Δ using (2).
6. Compute the CoG of $\overline{A}(T)$ and consider it as seed point for region growing method.
7. Initially, $\underline{A}(T) \leftarrow$ CoG.
8. The region $\underline{A}(T)$ grows by incorporating boundary region pixels by pixel assignment rule of (3).
9. Refine the boundary of lower approximation using the lower approximation refinement rule of (4).
10. Assign the lower approximation of T, $\underline{A}(T)$, to T.
11. Finally, apply closing morphology with a window of dimension 5×5 to smooth the surface of the tumor class.

3 Experimental Results and Discussions

This section presents the performance of the proposed brain tumor detection algorithm, along with a comparison with related methods. The proposed method uses robust rough-fuzzy c-means (rRFCM) [5] for segmentation of brain MR images. The methods compared are $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and BraTumIA [9]. The methods $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ use hard c-means (HCM), fuzzy c-means (FCM), and rough-fuzzy c-means (RFCM) [4], respectively, while mask generation, feature extraction and selection, and post-processing steps are same as those of the proposed method.

The value of fuzzifier is 2.00, while that of weight parameter for rough-fuzzy clustering algorithms, which represents relative importance of lower and boundary region is set to 0.51. The final cluster prototypes of HCM are used as the initial centroids of other clustering algorithms. Brain tumor image data used in this work were obtained from the MICCAI 2012 Challenge on Multimodal Brain Tumor Segmentation (http://www.imm.dtu.dk/projects/BRATS2012) organized by B. Menze A. Jakab, S. Bauer, M. Reyes, M. Prastawa, and K. Van Leemput. The challenge database contains fully anonymized images from the following institutions: ETH Zurich, University of Bern, University of Debrecen, and University of Utah. For a quantitative comparison of the performance of the proposed method with other methods, the ground truth of brain tumor for this data set is obtained from its corresponding website.

Based on the region of interest to be extracted in the output image and that in the ground truth, the false positive (FP), false negative (FN), true positive (TP), and true negative (TN) counts can be computed for each segmented image. The quantitative measures, namely, Dice coefficient and Youden index, are described as follows with the help of these counts:

- The Dice coefficient measures the overlap between the ground truth and the result, expressed as:

$$DC = \frac{2.\mathrm{TP}}{(\mathrm{FP}+\mathrm{TP})+(\mathrm{FN}+\mathrm{TP})}.$$

- The Youden index is defined as:

$$YI = \text{sensitivity} + \text{specificity} - 1,$$

$$\text{where sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \text{ and specificity} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}}.$$

Higher numbers of these metrics represent better overlapping in segmented image and ground truth image, indicating the significance of underlying algorithm. The metrics are calculated here for brain tumor that includes active tumor and edema. Figures 3 and 4 present the effectiveness of the proposed brain tumor segmentation algorithm for few images, while the heatmaps in Fig. 5 depict brain tumor segmentation results for fifty images quantitatively. From 3rd to 7th columns of Figs. 3 and 4, the TP, FP, and FN regions are represented in

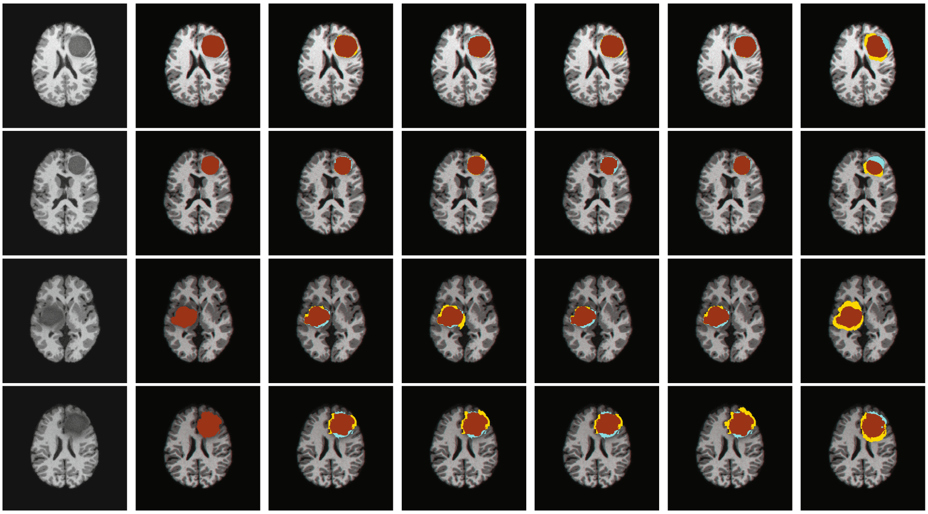


Fig. 3. Segmented images for simulated brain tumor images: input image, ground truth, proposed, $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and BraTumIA (from left to right) (Color figure online)

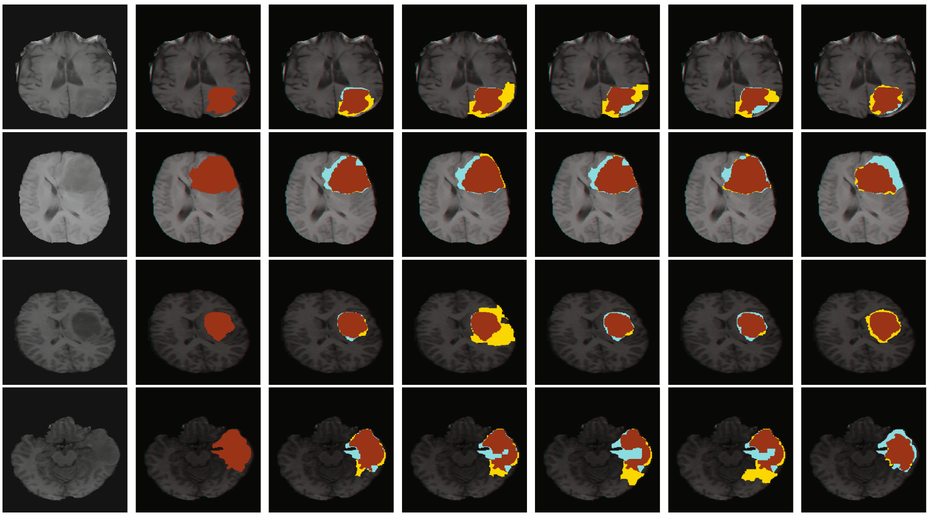
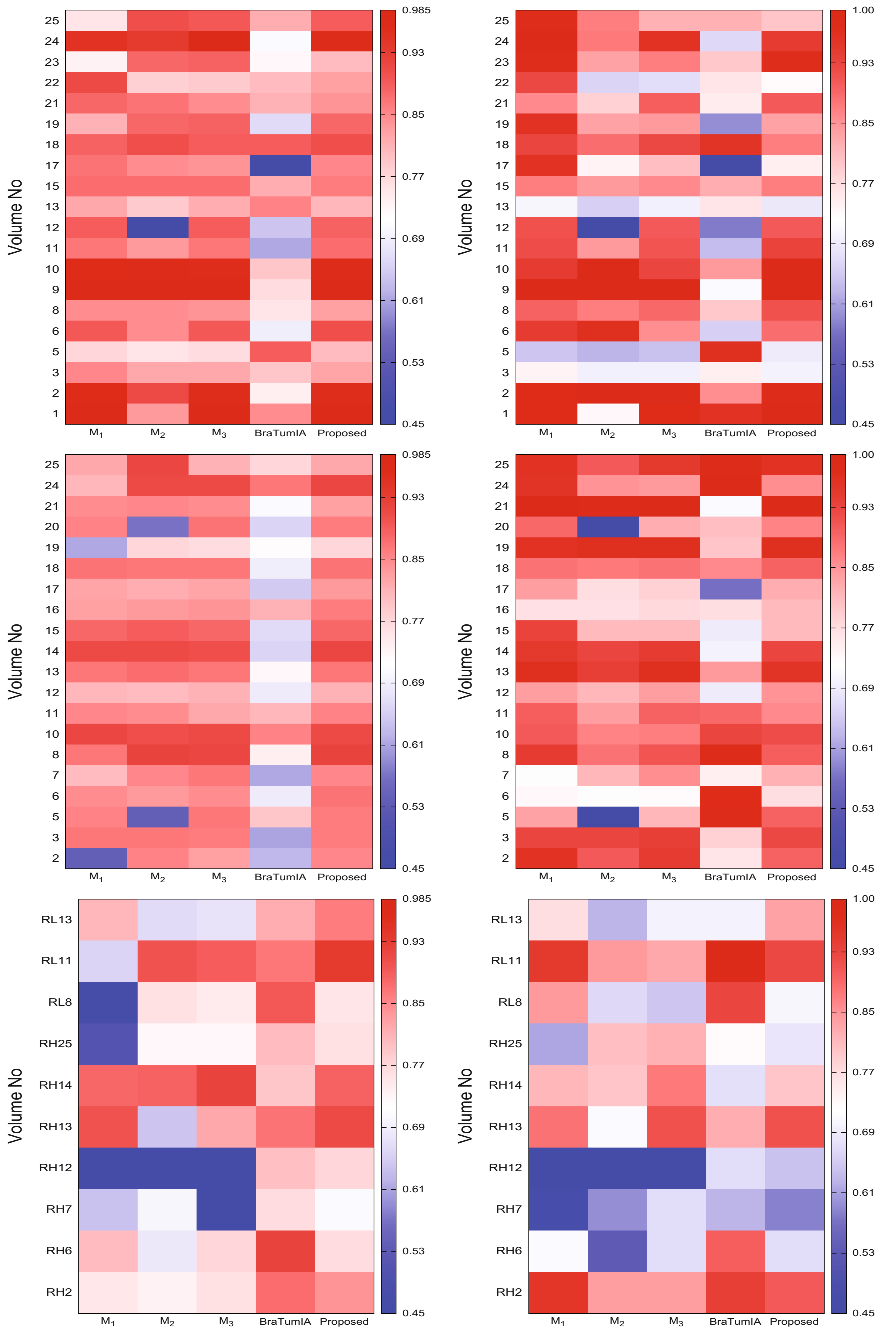


Fig. 4. Segmented images for real brain tumor images: input image, ground truth, proposed, $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$, and BraTumIA (from left to right) (Color figure online)



(a) Dice Coefficient (b) Youden Index

Fig. 5. Simulated high graded, simulated low graded, and real brain tumor

red, orange, and blue color, respectively, while red color in the images at second column represents region of interest to be segmented. In heatmaps for Dice coefficient and Youden index, the results below 0.45 are put into the bin numbered 1, while the results, starting from 0.45, lying within equally spaced intervals, are put into the corresponding bins.

3.1 Importance of Robust Rough-Fuzzy *C*-Means

The performance of the proposed method is extensively compared with that of the methods $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$. The qualitative results reported in columns 3rd to 6th of Figs. 3 and 4 compare the performance of the proposed method with that of the methods $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$, while Fig. 5 compares it quantitatively with respect to two quantitative indices on several brain tumor images. The proposed method attains better performance than $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ in 56, 73, 62 cases, respectively, out of 100 cases, irrespective of segmentation validation indices used. From the output images reported in Figs. 3 and 4, it can also be seen that there is a significant improvement in the segmentation results obtained using the proposed method as compared to other clustering methods.

From all these figures, it can also be concluded that the proposed post-processing method works well irrespective of clustering algorithms used. But, the clustering result actually forms the base for post-processing. So, more efficient is the clustering technique, more accurate the tumor segmentation. The best performance of the proposed method using rRFCM clustering algorithm is achieved due to the fact that the probabilistic membership function of the rRFCM handles efficiently overlapping partitions, while the possibilistic membership function of lower approximation of a cluster helps to discover arbitrary shaped cluster. In effect, good segmented regions are obtained using the proposed brain MR image segmentation algorithm.

3.2 Comparative Performance with Existing Method

Finally, the performance of the proposed method is compared with that of BraTumIA. The third and seventh columns of Figs. 3 and 4 represent the comparison between proposed segmentation results and the existing BraTumIA software tool qualitatively. It can be seen from Fig. 5 that the proposed method significantly gives better result than existing BraTumIA software tool. The proposed method attains higher values than BraTumIA in 42 and 30 cases with respect to Dice coefficient and Youden index, respectively, out of 50 cases each.

4 Conclusion

The main contribution of this paper is to introduce a post-processing step for brain tumor detection from MRI. The proposed post-processing method judiciously integrates the merits of morphological operations and the notion of rough set theory embedded into region growing technique. It improves the performance of the brain tumor segmentation method. Formulation of this method using rough set enables efficient handling of ambiguities at anatomical pathologic boundaries. The proposed post-processing method works significantly well, irrespective of clustering algorithms as shown by the experimental results. Hence, the proposed morphology and rough set based region growing post-processing method can be applied to any clustering based brain tumor segmentation approach. Several quantitative measures are used to evaluate the performance of the proposed brain tumor segmentation method. Finally, the effectiveness of the proposed method is demonstrated both qualitatively and quantitatively, along with a comparison with other related algorithms, on a set of synthetic and real brain MR images.

Acknowledgments. This work is partially supported by the Indian National Science Academy, New Delhi, India (grant no. SP/YSP/68/2012).

References

1. Fletcher-Heath, L.M., Hall, L.O., Goldgof, D.B., Murtagh, F.R.: Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. Artif. Intell. Med. **21**, 43–63 (2001)
2. Hsieh, T.M., Liu, Y.-M., Liao, C.-C., Xiao, F., Chiang, I.-J., Wong, J.-M.: Automatic segmentation of meningioma from non-contrasted brain MRI integrating fuzzy clustering and region growing. BMC Med. Inform. Decis. Mak. **11**, 54 (2011)
3. Iscan, Z., Dokur, Z., Ölmez, T.: Tumor detection by using zernike moments on segmented magnetic resonance brain images. Expert Syst. Appl. **37**(3), 2540–2549 (2010)
4. Maji, P., Pal, S.K.: Rough set based generalized fuzzy C-Means algorithm and quantitative indices. IEEE Trans. Syst. Man Cybern. Part B: Cybern. **37**(6), 1529–1540 (2007)
5. Maji, P., Paul, S.: Rough-fuzzy clustering for grouping functionally similar genes from microarray data. IEEE/ACM Trans. Comput. Biol. Bioinf. **10**(2), 286–299 (2013)
6. Maji, P., Roy, S.: Rough-Fuzzy Clustering and Unsupervised Feature Selection for Wavelet Based MR Image Segmentation. PLoS ONE **10**(4), e0123677 (2015). doi:10.1371/journal.pone.0123677
7. Menze, B.H., Leemput, K., Lashkari, D., Weber, M.-A., Ayache, N., Golland, P.: A generative model for brain tumor segmentation in multi-modal images. Med. Image Comput. Comput.-Assist. Interv. **6362**, 151–159 (2010)
8. Nie, J., Xue, Z., Liu, T., Young, G.S., Setayesh, K., Guo, L., Wong, S.T.: Automated brain tumor segmentation using spatial accuracy-weighted hidden markov random field. Comput. Med. Imaging Graph. **33**(6), 431–441 (2009)
9. Porz, N., Bauer, S., Pica, A., Schucht, P., Beck, J., Verma, R.K., Slotboom, J., Reyes, M., Wiest, R.: Multi-modal Glioblastoma segmentation: man versus machine. PLoS ONE **9**(5), e96873 (2014). doi:10.1371/journal.pone.0096873
10. Stadlbauer, A., Moser, E., Gruber, S., Buslei, R., Nimsky, C., Fahlbusch, R., Ganslandt, O.: Improved delineation of brain tumors: an automated method for segmentation based on pathologic changes of H-MRSI metabolites in gliomas. NeuroImage **23**(2), 454–461 (2004)

11. Sung, Y.-C., Han, K.-S., Song, C.-J., Noh, S.-M., Park, J.-W.: Threshold estimation for region segmentation on MR image of brain having the partial volume artifact. In: Proceeding of the 5th International Conference on Signal Processing, vol. 2, pp. 1000–1009 (2000)
12. Xie, K., Yang, J., Zhang, Z.G., Zhu, Y.M.: Semi-automated brain tumor and edema segmentation Using MRI. Eur. J. Radiol. **56**(1), 12–19 (2005)