Brain Tumor Segmentation of Normal and Pathological Tissues Using K-mean Clustering with Fuzzy C-mean Clustering

Ravi Shanker(✉) and Mahua Bhattacharya

ABV-IIITM, Gwalior, Madhya Pradesh, India
rsmiet60@gmail.com, mb@iiitm.ac.in

Abstract. Segmentation of brain tumor from magnetic resonance imaging is a time consuming and critical task due to unpredictable characteristics of tumor tissues. In this paper, we propose a new tissue segmentation algorithm that segments brain MR images into gray matter (GM), white matter (WM), cerebrospinal fluid (CSF), tumor and edema. It is crucial to segment the normal and pathological tissues simultaneously for treatment planning. K-mean clustering algorithm has minimal computation time, and fuzzy c mean clustering has advantages in the aspect of accuracy on the soft tissues. So we are integrating the K-mean clustering algorithm with Fuzzy C-means clustering algorithm for segmenting the brain magnetic resonance imaging. First, we segment the abnormal region from T_2-weighted FLAIR modality based on k mean clustering algorithm integrated with fuzzy c mean algorithm. And in the next stage, we segment the tumor from T_1-weighted contrast enhancement modality T_{1ce}. We used T_1, T_{1c} , T_2 and flair images of 60 subject suffering from high graded and low grade glioma, and 20 T_1-weighted anatomical models of normal brains.

Keywords: Medical image segmentation · Brain magnetic resonance · K-mean clustering · Fuzzy C-mean clustering

1 Introduction

Brain MR image segmentation is a very challenging and important task that is required for diagnosing neurological diseases and brain tumors. Brain tumors have specialties such as image size, shape, location, and image intensities. Brain magnetic resonance imaging comprises of healthy tissues (gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF)) and pathological tissues (tumor, edema and necrosis) [1,2]. If there are tumor and edema, intensity characteristics of the nearby tissues change and it may deform the neighbouring structures of the gray matter, white matter and cerebrospinal fluid. Glioma has the glial cells which have a high mortality rate, and it is due to glial cells of the brain, and it shows fast growth into the normal tissues. Segmentation of 1500-200 images of 512×512 size takes about 2–4 h with 14%–22% differences.

J.M.R.S. Tavares and R.M. Natal Jorge (eds.), *VipIMAGE 2017*,
Lecture Notes in Computational Vision and Biomechanics 27,
DOI 10.1007/978-3-319-68195-5_31

Segmentation of pathological tissues such as edema and tumor is quite tough task due to complex shape, unclear boundaries and intensity distribution [3].

In recent years, there has been growing interest in the segmentation of the tumor and edema from MRI images. For segmentation of gliomas in 3D MR multimodality medical imaging Discriminative random decision forest framework and tissue-specific Gaussian mixture models are used for voxel-wise probabilistic classification. They segmented the tumor, edema, necrotic core and active cells of brain [4,5]. Peodia et al. [6] proposed a hybrid method for brain tumor segmentation using the graph cut segmentation method and competitive expectation maximization algorithm.

There is the study of image segmentation technique based on hierarchical clustering, k-mediods and k-means clustering algorithms. K-mean clustering algorithm is easy to implement and better performance than hierarchical clustering and k-mediods clustering algorithms. The k-means clustering algorithm is simple and fast to run of big datasets, but it is insufficient to detect the pathological brain tissues such as tumor, edema and necrosis completely [7]. Fuzzy C-means clustering algorithm has more information of the original image to detect pathological cells accurately compared to the K-means. But FCM clustering algorithm takes long execution time [8,9].

Single MRI modality is not enough to segment the normal and abnormal regions because of the appearance and nature of brain tumors. Multimodality MRI sequences are used for delineation and diagnosis of pathological subregions. Multimodality MRI sequence comprises of T_1-weighted MRI (T_1), T_1-weighted MRI with contrast enhancement (T_{1ce}), T_2-weighted MRI and T_2-weighted MRI with fluid-attenuated inversion recovery (T_{2FLAIR}) or FLAIR. T_1 -weighted modality allows for an easy annotation of GM, WM and CSF. Tumor borders show brighter in the T1-weighted contrast-enhanced images [10]. Edema region appears brighter and it surrounds the tumor region in the T_2 weighted MRI. T_{2FLAIR} MRI modality separates the cerebrospinal fluid to edema region [11].

We are integrating the k mean clustering algorithm to fuzzy c mean clustering algorithm for segmentation of the normal and abnormal tissues in brain MR images. The novelty of the proposed work is that first, we are segmenting the edema from the FLAIR modality and tumor from the t_{1ce} modality. And then normalized summation of the tumor and edema region. So we are using the multimodality MR imaging for their different features. The tumor and edema region represents the different gray level to the normal tissues. The Jaccard coefficient or Dice similarity coefficient are used to quantitatively Calculate the segmented results to the overlapping with the ground truth results. There are the three different validation techniques: mutual information, Dice similarity coefficient and receiver operating characteristic [12]. Mutual information should be used for the spatial alignment evaluation, and the dice coefficient is the best for quantitative evaluation. In the lack of ground truth segmentation results with the brain tumor data base, so researchers evaluate their algorithms on the less number of images. It makes the difficult to correlate the evaluation of the different algorithms on the brain MRI images [14]. Data set has been taken from the Brats challenge at MICCAI2013 and brainweb [15,16].

2 Methodology

In the methodology, we give the details about the materials, brain magnetic resonance imagining data sets, and the methods used to perform brain MR normal and abnormal tissue segmentation algorithm. Flow diagram of overall algorithm is given in the Fig. 1. There are the 6 steps in the algorithm and it will discussed in the following sections.

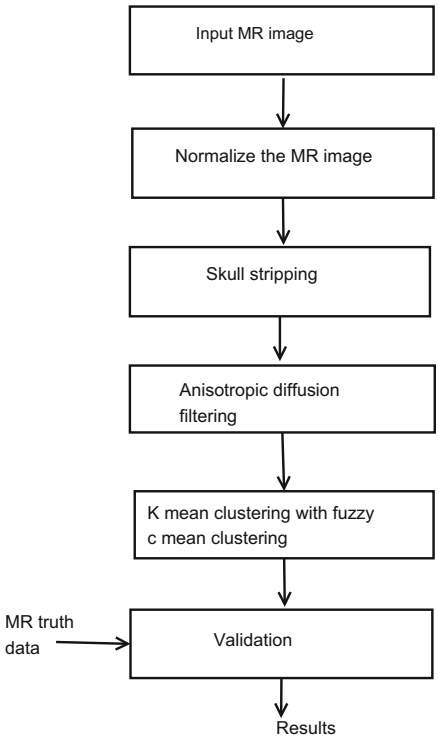

Fig. 1. Flow chart of brain normal and abnormal tissues segmentation algorithm

2.1 Brain MR Images

We worked with a data set of 80 patients. BRATS2013 data set, which includes 60 multimodal MRI images from patients with high- and low-grade gliomas. The dataset comprise of the T1, T1c, T2 and flair MR images. One volume of the brats multimodal medical imaging contain 155 slices. The truth data set of abnormal tissues are given in the data set. The data set of Brain web comprises of 20 T1 weighted anatomical models of normal brains. The given dataset multimodal medical imaging contain the 155 MRI slices. The slices of the flair modality are given into the Fig. 2.

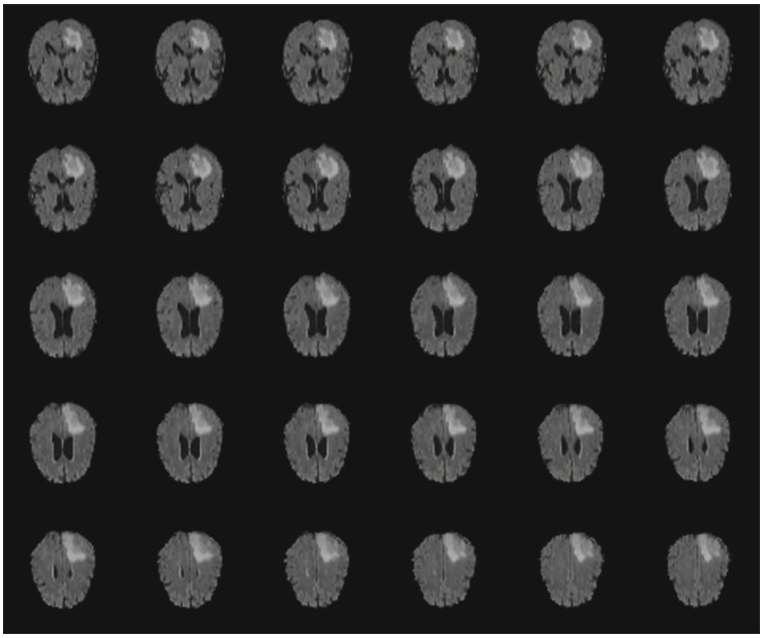

Fig. 2. Axial slices of T_2 weighted fluid-attenuated inversion recovery from slice 80 to slice 100

Unimodal medical imaging methods are easy to implement as compared to multimodality medical imaging. We can not segment the full tumor region from the single MRI modality. Necrosis and active tumor region were segmented from the T1-weighted with contrast enhancement, where as edema region was segmented based one the T_{2FLAIR}. In the Fig. 3 one axial slice of the T_1, T_{1ce}, T_2 and T_{2FLAIR} modalities are shown.

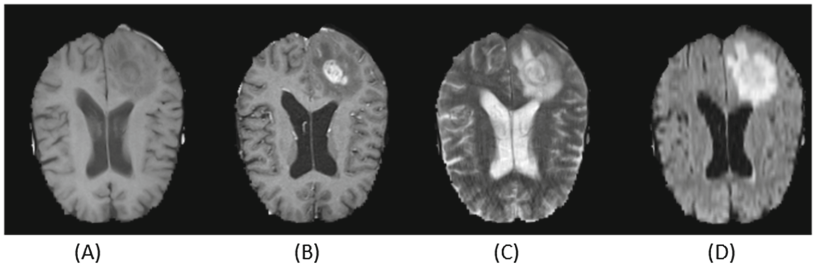

Fig. 3. One axial slice of MR image, (A) T_1, (B) T_{1ce}, (C) T_2, (D) T_{2FLAIR} modality

2.2 Normalization of the MR Images

We normalize the intensity range of the images to [0 255] range. It is necessary for quantitative texture analysis. Due to the inter scan and intra scan image intensity variation we normalize the image intensity. There are the MRI intensity normalization methods: histogram equalization, contrast stretch normalization, histogram stretching, gaussian kernel normalization and intensity scaling histogram normalization. The histogram normalization method give the best results as compared to other normalization methods [13]. The mathematical expression of MRI intensity normalization is given below.

$$s(i,j) = \frac{x(i,j) - x_{min}}{x_{max} - x_{min}} \times (NEW_{max} - NEW_{min}) + NEW_{min} \tag{1}$$

where x(i, j) is the image pixel, x_{min} and x_{max} is the smallest and largest gray level in the image. NEW_{max} and NEW_{min} is the new maximum and new minimum intensity levels in the image.

2.3 Skull Stripping and De-noising

Skull stripping is an important step to remove the non-cerebral tissues such as connective tissues, muscle, skull, fat and skin, which are not region of interest. Skull stripping is done by the Brain surface extractor(BSE) algorithm. BSE algoritihm only extract the skull and other irregularities of MRI images. Here we are using the medical image processing, analysis, and visualization tool box for brain surface algorithm [14]. There are the four steps in brain surface extractor algorithm. In the first step filtering the image to remove irregularities, in the second step detecting edges in the image, in the third step performing morphological erosions and brain isolation and in the fourth step performing surface cleanup and image masking. The noise is removed from the gray level MRI image. The median filter is used for removing the noise and high-frequency components.

2.4 K-mean Clustering with Fuzzy C-mean Clustering

K-mean clustering algorithm is simple and fast on big data sets but it does not segment the tumor effectively mainly if it is malignant. So first we select the k = 4 and calculate the centriods of the GM, WM, CSF, tumor and edema. Centriods are calculated in we keep it into the cluster centroids vector V. and this centriod vector are given into the FCM algorithm. FCM is the power full and best known method that has been extensively used in MRI, where pixels are partially classified for each class. If U is the fuzzy membership, and V is the cluster centroids vector, then $J_m(U, V)$ is the objective function of Fuzzy C-mean algorithm.

$$J_m(U,V) = \sum_{j=1}^{N} \sum_{i=1}^{C} \mathbf{u}_{ij}^{m} (\mathbf{x}_j, \mathbf{v}_i) \tag{2}$$

In the above equation, m = 2 and $X = x_1, x_2, x_3, x_4.....x_N$ is a $p \times N$ data matrix, where N represents the number of feature vectors (pixel numbers in the image) and p represents the dimension of each x_j feature vector, and C is the number of clusters in the Fuzzy c-mean clustering algorithm. $u_{ij} \subseteq U(p \times N \times C)$ is the membership function of vector x_j to the i^{th} cluster, which satisfies $u_{ij} \in [0, 1]$ and $\sum_{i=1}^{C} u_{ij} = 1$ $(j = 1, 2.....N)$. u_{ij} is the membership function.

$$u_{ij} = \frac{1}{\sum_{k=1}^{C} \left(\frac{d(x_j, v_i)}{d(x_j, v_k)} \right)^{2/(m-1)}} \tag{3}$$

where $V = v_1, v_2.....v_c$, which is a $p \times C$ matrix shows feature center of the cluster .

$$v_i = \frac{\sum_{j=1}^{N} (u_{ij})^m x_j}{\sum_{j=1}^{N} (u_{ij})^m} \tag{4}$$

$m = 2$ controls the degree of fuzzines and it is weighting exponent on each fuzzy membership. $d^2(x_j, v_i)$ is a measurement of similarity between x_j and v_i

$$d^2(x_j, v_i) = \|x_j - v_i\|^2 \tag{5}$$

The pixel intensity is represented by the feature vector X in MRI, so keep p = 1. The FCM algorithm optimize the cluster iteratively $J_m(U, V)$ and it continuously updates the U and V with the continuous update of U and V, until it hold the condition $|U^{l+1} - U^l| \leq \varepsilon$ where l is the number of iterations in the FCM clustering algorithm.

2.5 Validation

The validation of the results are shown on the basis of Dice similarity coefficient, sensitivity, Specificity. The mathematical expression of Dice similarity indexes, sensitivity, Specificity is given as,

$$Dice = \frac{2 \times TP}{2 \times TP + FP + FN} \times 100 \tag{6}$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100 \tag{7}$$

$$Specificity = \frac{TN}{TN + FP} \times 100 \tag{8}$$

Where True positive (TP) correctly classified images having brain tumor, true negative (TN) correctly classified images don't have the brain tumor and false negative (FN) incorrectly classified images have tumor but not detected the tumor. Dice measures the overlapping between segmented results and ground truth.

3 Results and Discussion

3.1 Segmentation of GM, WM and CSF

The results of proposed algorithm are evaluated on the Brainweb data set. The data set contain the 20 T_1-weighted anatomical models of normal brains, the specific parameters: SFLASH sequence with TE = 9.2 ms, TR = 22 ms, flip angle = 30° and 1 mm isotropic voxel size. T_1-weighted MRI modality is used for segmentation of GM, WM and CSF. Brainweb data set comprises of truth data of the soft tissues and hard tissues. In the Fig. 4, there are the three rows, and each row contains T_1-weighted MRI modality. We demonstrate the results of our algorithm in the form of images and quantitative way. The description is given as: (A) shows the single slice of the T_1 weighted MRI modality, (B) skull stripped image in this step extract the skull and other irregularities from the images, segmentation of GM, WM and CSF fluid is shown in the (C), (D) and (E) columns. Gray matter does not mean that it is the gray color it is pink color in our brain. The quantitative segmentation results of GM, WM and CSF are shown in the Table 1. The average Dice similarity indexes are 93.20% for gray matter, 92.91% for white matter and 94.06% for the 97.09% for cerebropinal fluid. The sensitivity and specificity of the GM, WM and CSF are above 90%.

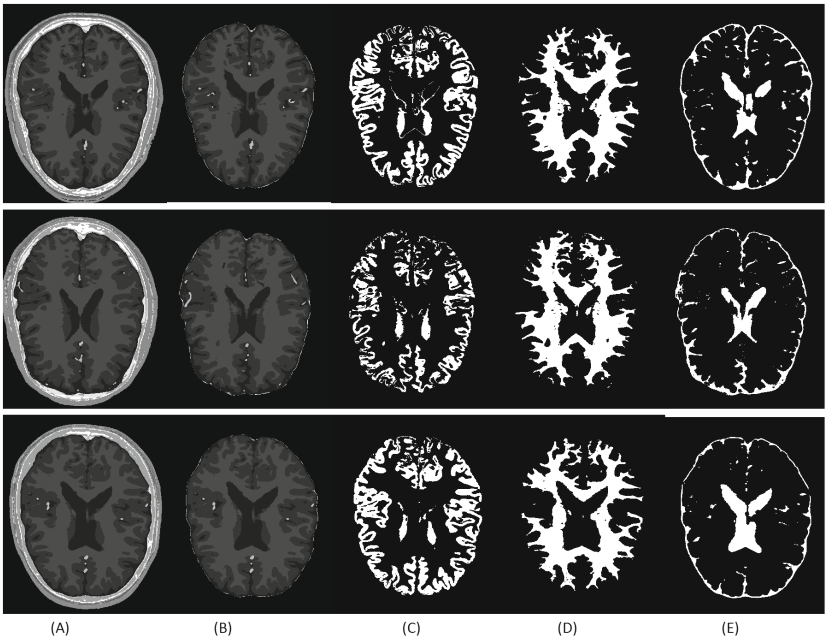

Fig. 4. One axial slice of different T_1 weighted MR images from first row to third row: (A) Original MR image. (B) Skull stripped image (C) Gray matter (D) White matter (E) Cerebrospinal fluid

Table 1. Measure of overlap for segmented normal tissues (GM, WM and CSF in (%)) of Brainweb data set

Patient	GM			WM			CSF		
	Dice	Sens	Spec	Dice	Sens	Spec	Dice	Sens	Spec
1	89.05	82.06	95.31	92.24	88.11	97.25	93.03	86.97	96.74
2	97.62	86.40	99.53	90.53	86.60	97.65	98.00	97.50	98.00
3	95.42	98.00	96.42	91.40	96.72	95.12	99.60	99.00	99.50
4	96.60	99.80	97.13	94.60	91.40	97.72	97.09	98.40	98.20
5	88.90	82.01	94.80	92.24	87.11	97.25	93.14	87.16	97.75
6	94.62	99.00	98.15	95.20	96.80	99.00	88.70	98.40	98.00
7	90.24	84.62	89.62	94.16	88.77	97.44	88.85	79.99	97.09
Average	93.20	90.27	95.85	92.91	90.79	97.27	94.06	92.59	97.09

3.2 Segmentation of Tumor and Edema

The Brats data set consist of the 60 patients data set. All data set contain four modalities i.e. T_1, T_{1ce}, T_2 and T_{2FLAIR}. In our study, we have taken only T_{2FLAIR} or FLAIR and T_{1ce} modalities. In the Fig. 5 the axial slice of FLAIR modality in the first column, FLAIR contain the edema and in the second

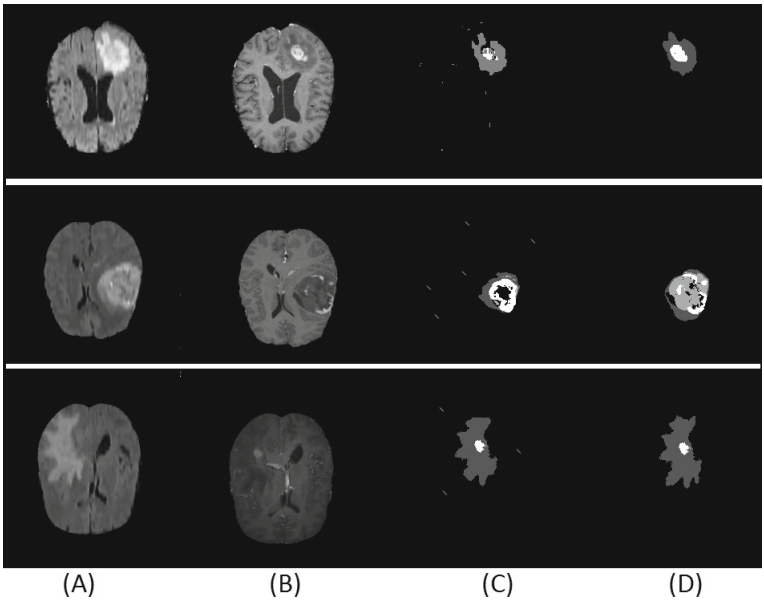

Fig. 5. One axial slice of: (A) T_{2FLAIR}, (B) T_{1ce}t, (C) segmented tumor and edema (edema in gray and tumor in white color), (D) ground truth MRI image, from first row to third row

column T_{1ce} which contain the information of tumor. The segmented results of the edema and tumor in the third row, the edema is the gray color and tumor is in the white color. The quantitative results on the Brats2013 data set are given in the Table 2. The average dice similarity indexes are 76.65% for tumor and 71.17% for edema. The average sensitivity and specificity for tumor and edema are above 69%.

Table 2. Measure of overlap for segmented abnormal tissues (tumor and edema in (%)) of Brats2013 data set

Patient	Tumor			Edema		
	Dice	Sens	Spec	Dice	Sens	Spec
1	52.61	50.26	91.50	36.84	34.50	89.15
2	80.20	65.92	98.62	76.11	68.90	97.65
3	64.32	76.30	87.50	83.06	80.60	98.60
4	95.28	91.70	98.20	96.06	99.65	97.20
5	66.21	58.62	95.12	63.64	67.22	89.65
6	96.33	94.62	97.76	71.40	56.00	98.46
7	82.11	70.60	98.60	79.45	78.40	95.00
Average	76.65	72.57	95.32	71.17	69.32	95.10

3.3 Comparison Analysis

Table 3 represents the results of proposed method and the methods in [4–6] respectively. The mean dice similarity index coefficient in Table 3 demonstrate a very competitive performance of the proposed method. The average dice coefficient value obtained by the k-mean clustering algorithm integrated with fuzzy c-mean clustering algorithm is 76.65% for tumor and 71.11% for edema. The standard deviation of dice coefficient for tumor is 16.40% and for edema is 18.61%. The overall dice coefficient for tumor and edema is $76.65 \pm 16.40\%$ and $71.11 \pm 18.61\%$ respectively.

Table 3. Quantitative evaluation of dice similarity index of tumor and edema

Authors	Tumor	Edema
Geremia et al. [4]	68 ± 18	56 ± 17
Pedoia et al. [5]	70 ± 21	59 ± 18
Zikic et al. [6]	71 ± 24	70 ± 09
Proposed method	76.65 ± 16.40	71.11 ± 18.61

Quantitative evaluation of k-mean, fuzzy c- mean and k-mean integrated with fuzzy c-mean clustering algorithms are demonstrated in Table 4. The run time

of KM, FCM and KFCM are 2.90 s, 18.62 s and 8.92 s. The execution time of KFCM is less as compared to KM and FCM. Dice similarity index, sensitivity and specificity of the KFCM algorithm better than the KM and FCM clustering algorithms.

Table 4. The performance matrices of KM, FCM and KFCM clustering algorithms

Type of cluster	Time(s)	Tumor			Edema		
		Dice	Sens	Spec	Dice	Sens	Spec
KM	2.90	68.72	64.35	90.37	65.25	64.67	91.67
FCM	18.62	73.27	70.62	94.37	69.24	66.46	94.20
KFCM	8.92	76.65	72.57	95.32	71.17	69.32	95.10

4 Conclusion

In this study, we segmented brain MR images into normal tissues such as gray matter, white matter, and cerebrospinal fluid along with the abnormal tissues such as edema and tumor. The data of 60 patients suffering from the high graded glioma and low graded glioma are used along with 20 T_1-weighted anatomical models normal brains. We registered T_1-weighted contrast enhancement and T_2-weighted FLAIR MR images into single co-ordinate system. Multimodality MR images are the most effective for treatment and diagnosis of brain tumor. We integrated the k-mean clustering algorithm with fuzzy c mean clustering to segment the brain tumor accurately and in minimal execution time. For evaluation, we used the specificity, sensitivity and Dice similarity index. The results showed that average dice similarity indexes are 93.20% for GM, 92.91% for WM, 94.06% for CSF on the Brain web data set and 76.65% for tumor, and 69.32% for edema on the Brats2013 data set.

References

1. Demirhan, A., Törü, M., Güler, İ.: Segmentation of tumor and edema along with healthy tissues of brain using wavelets and neural networks. IEEE J. Biomed. Health Inform. **19**(4), 1451–1458 (2015)
2. Demirhan, A., Guler, I.: Image segmentation using self-organizing maps and gray level co-occurrence matrices. J. Fac. Eng. Arch. Gazi Univ. **25**(2), 285–291 (2010)
3. Abdel-Maksoud, E., Elmogy, M., Al-Awadi, R.: Brain tumor segmentation based on a hybrid clustering technique. Egypt. Inform. J. **16**(1), 71–81 (2015)
4. Geremia, E., Menze, B.H., Ayache, N.: Spatial decision forests for glioma segmentation in multi-channel MR images. MICCAI Chall. Multimodal Brain Tumor Segmentation **34** (2012)
5. Zikic, D., Glocker, B., Konukoglu, E., Criminisi, A., Demiralp, C., Shotton, J., Thomas, O.M., Das, T., Jena, R., Price, S.J.: Decision forests for tissue-specificsegmentation of high-grade gliomas in multi-channel MR. In: Medical Image Computing and Computer-Assisted Intervention MICCAI, pp. 369–376 (2012)

6. Pedoia, V., Balbi, S., Binaghi, E.: Fully automatic brain tumor segmentation by using competitive EM and graph cut. In: International Conference on Image Analysis and Processing, pp. 568–578. Springer International Publishing (2015)
7. Deng, W.-Q., Li, X.-M., Gao, X., Zhang, C.-M.: A modified fuzzy C-means algorithm for brain MR image segmentation and bias field correction. J. Comput. Sci. Technol. **31**(3), 501–511 (2016)
8. Pham, D.L., Prince, J.L.: An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities. Pattern Recogn. Lett. **20**(1), 57–68 (1999)
9. Kaus, M., Warfield, S.K., Jolesz, F.A., Kikinis, R.: Adaptive template moderated brain tumor segmentation in MRI. In: Proceedings of Bildverarbeitung fur die Medizin, pp. 102–106 (1999)
10. Song, Y., Ji, Z., Sun, Q., Zheng, Y.: A novel brain tumor segmentation from multimodality MRI via a level-set-based model. J. Signal Process. Syst.,1–9 (2016)
11. Anitha, V., Murugavalli, S.: Brain tumour classification using two-tier classifier with adaptive segmentation technique. IET Comput. Vis. **10**(1), 9–17 (2016)
12. Kaya, I.E., Pehlivanl, A.Ç., Sekizkardeş, E.G., Ibrikci, T.: PCA based clustering for brain tumor segmentation of T1w MRI images. Comput. Methods Progr. Biomed. **140**, 19–28 (2017)
13. Loizou, C.P., Pantziaris, M., Seimenis, I., Pattichis, C.S.: Brain MR image normalization in texture analysis of multiple sclerosis. In: IEEE International Conference on Information Technology and Applications in Biomedicine, pp. 1–5 (2009)
14. NIH Center for information technology. https://mipav.cit.nih.gov/
15. Brats multimodal brain tumor segmentation (2013). http://martinos.org/qtim/miccai2013/data.html
16. Brain web simulated brain database. http://brainweb.bic.mni.mcgill.ca/brainweb/