

Automatic Brain Tumor Segmentation in Multispectral MRI Volumetric Records

László Szilágyi^{1,2,3}(✉), László Lefkovits², Barna Iantovics⁴, David Iclănzan²,
and Balázs Benyó¹

¹ Department of Control Engineering and Information Technology,
Budapest University of Technology and Economics, Budapest, Hungary
`lszilagyi@iit.bme.hu`

² Sapientia University of Transylvania, Tîrgu-Mureş, Romania

³ Canterbury University of Christchurch, Christchurch, New Zealand

⁴ Department of Informatics, Petru Maior University of Tîrgu-Mureş,
Tîrgu-Mureş, Romania

Abstract. The aim of this study was to establish a multi-stage fuzzy *c*-means (FCM) framework for the automatic and accurate detection of brain tumors from multimodal 3D magnetic resonance image data. The proposed algorithm uses prior information at two points of the execution: (1) the clusters of voxels produced by FCM are classified as possibly tumorous and non-tumorous based on data extracted from train volumes; (2) the choice of FCM parameters (e.g. number of clusters, fuzzy exponent) is supported by train data as well. FCM is applied in two stages: the first stage eliminates the most part of non-tumorous tissues from further processing, while the second stage is intended to accurately extract the tumor tissue clusters. The algorithm was tested on 13 selected volumes from the BRATS 2012 database. The achieved accuracy is generally characterized by a Dice score in the range of 0.7 to 0.9. Tests have revealed that increasing the size of the train data set slightly improves the overall accuracy.

Keywords: Image segmentation · Fuzzy clustering · Tumor detection · Magnetic resonance imaging

1 Introduction

The early detection of brain tumors is utmost important as it can save human lives. The accurate segmentation of brain tumors is also essential, as it can assist the medical staff in the planning of treatment and intervention. The manual segmentation of tumors requires plenty of time even for a well-trained expert.

Research supported by the Hungarian National Research Funds (OTKA), Project no. PD103921. The work of L. Lefkovits was supported by The Sectorial Operational Program Human Resources Development POSDRU/159/1.5/S/137516 financed by the European Social Found and by the Romanian Government.

A fully automated segmentation and quantitative analysis of tumors is thus a highly beneficial service. However, it is also a very challenging one, because of the high variety of anatomical structures and low contrast of current imaging techniques which make the difference between normal regions and the tumor hardly recognizable for the human eye [1]. Recent solutions, usually assisted by the use of prior information, employ various image processing and pattern recognition methodologies like: combining multi-atlas based segmentation with non-parametric intensity analysis [2], AdaBoost classifier [3], level sets [4], active contour model [5], graph cut distribution matching [6], diffusion and perfusion metrics [7], confidence guided discriminative classifier [8], 3D blob detection [9], and support vector machine [10].

The main goal of our research work is to build a reliable procedure for brain tumor detection from multimodal MRI records, based on semi-supervised clustering algorithms, using the MICCAI BRATS data set that contains several dozens of image volumes together with ground truth provided by human experts. As a first step, in the current paper we present preliminary results achieved using a two-stage fuzzy *c*-means clustering cascade algorithm and 13 selected volumes from the above mentioned data set.

2 Materials and Methods

2.1 BRATS Data Sets

Brain tumor image data used in this work were obtained from the MICCAI 2012 Challenge on Multimodal Brain Tumor Segmentation [11]. The challenge database contains fully anonymized images originating from the following institutions: ETH Zürich, University of Bern, University of Debrecen, and University of Utah. The image database consists of multi-contrast MR scans of 30 glioma patient, out of which 20 have been acquired from high-grade (anaplastic astrocytomas and glioblastoma multiforme tumors) and 10 from low-grade (histological diagnosis: astrocytomas or oligoastrocytomas) glioma patients. For each patient, multimodal (T1, T2, FLAIR, and post-Gadolinium T1) MR images are available. All volumes were linearly co-registered to the T1 contrast image, skull stripped, and interpolated to 1 mm isotropic resolution. All images are stored as signed 16-bit integers, but only positives values are used. Each image set has a truth image which contains the expert annotations for “active tumor” and “edema”.

In our application, each voxel in a volume is represented by a four-dimensional feature vector:

$$\mathbf{x} = [\log(x^{(T1)}), \log(x^{(T2)}), \log(x^{(T1C)}), \log(x^{(FLAIR)})]^T.$$

Those voxels which have zero intensity in any of the channels are neglected. Most of these are voxels outside the volume of interest, but there are a few others as well. The intensity information of voxels of a whole volume are collected in a large matrix whose number of rows is 4, while the number of columns is the number of actual voxels, somewhere between 1 and 2 million. This matrix will

represent the input data for the FCM cascade. We also store the position of each voxel, so that we are able to localize them after clustering. The log values of intensities are likely to be in the range between 3 and 8.

2.2 The FCM Cascade

The FCM cascade algorithm has the main goal of effectively separating homogeneous areas in the volume. Further on, based on decision support built upon prior information, it also distinguishes tumor tissues from normal brain tissues.

The conventional FCM algorithm provides the input data set a fuzzy partitioning into a previously set number (c) of clusters, based on the minimization of a quadratic objective function that contains a parameter referred to as fuzzy exponent. This parameter $m > 1$ regulates the fuzzyness of the partition. The smaller the value of the fuzzy exponent is, the closer the partition will be to the crisp one [12]. FCM usually places the cluster prototypes in areas where lots of input vectors are in the neighborhood. In order to provide an accurate clustering, we should know in advance the exact number of such accumulation points in the 4D color space, and be able to initialize a cluster prototype in the proximity of each. Obviously this is not the case. Even if we could do this somehow, clustering would represent a huge computational burden because of the hundreds of clusters and millions of input vectors. To avoid this case, we propose to perform FCM in two stages. The first stage will help us get rid of the most part of the input data, those vectors which are far from all tumor tissue intensities that we have stored in a predefined atlas. In the second stage we only have voxels of intensities that are close to those of tumor patterns. Cluster prototypes obtained in the second stage are individually analyzed, to separate tumor clusters from normal ones.

So in the first stage, fuzzy c -means is applied to the whole set of n voxels in the volume. The number of clusters c is typically varied between 6 and 20. When this first FCM stage is ready, the voxels are categorized into c clusters, each represented by its cluster prototype \mathbf{v}_i , $i = 1 \dots c$. These cluster prototypes are individually investigated by the decision support system which decides whether they are suspected of containing tumor tissues or not. Those clusters whose centroid vector is distant from tumor intensities, are excluded from the further stages. The decision support at this stage usually keeps up to 3 of the clusters obtained in the first FCM. The remaining n' voxels, or more precisely, the collection of feature vectors that represent them, will serve as input data for the second stage of the FCM cascade. Fuzzy c -means clustering is applied again, this time with c' clusters, also varied between 6 and 20. Final cluster prototypes are checked again by the decision support system. This time all clusters having their prototypes in the proximity of tumor intensities will be declared tumor tissues (positive), while all others negative. The fuzzy exponent may vary from stage to stage within the interval $[1.5, 2.0]$, that is why we denoted by m and m' the exponent applied in the first and second stage, respectively.

2.3 FCM Initialization

Especially in multi-dimensional problem, the FCM algorithm is highly sensitive to prototype initialization. Basically it is advised to attempt placing the initial cluster prototypes far from each other, possibly in accumulation areas of input vectors. Randomly chosen input vectors can produce high variability among different runs. So a deterministic rule is needed for a stable solution.

Our algorithm applies FCM clustering in each dimension d on the scalar data $\log(x_1^{(d)}), \log(x_2^{(d)}) \dots \log(x_n^{(d)})$ to produce three clusters, whose cluster prototypes $\nu_1^{(d)}, \nu_2^{(d)}$, and $\nu_3^{(d)}$ will serve as grid points in the given dimension d . The 4D grid formed this way defines $3^4 = 81$ grid points, which are all treated as potential cluster seeds for the 4D clustering problem. Let us denote these potential seeds by $\mathbf{w}_i, i = 1 \dots 81$. Out of these, we choose initial cluster prototypes those ones, which have lowest average square distance from input vectors, computed with the formula $\Delta(\mathbf{w}_i) = \sum_{k=1}^n \|\mathbf{x}_k - \mathbf{w}_i\|^2$.

2.4 Decision Support

After the first stage of the FCM cascade, it is necessary to separate those clusters that are likely to contain tumor tissues from the other clusters. After the second stage of the cascade, a decision has to be taken, which are the tumor tissue clusters and which are declared negative. These decisions rely on prior information extracted from train volumes. A series of previously performed FCM runs have established clusters that were declared tumor tissues or negative ones based on the available ground truth. In the testing phase, a k -nearest neighbors algorithm is employed to decide, whether the extracted clusters contain tumor tissues or not. Ground truth information extracted from the test volume is never used by the decision support algorithm.

2.5 Evaluation of Accuracy

The Jaccard index (JI) is a normalized score of accuracy, computed as $JI = \frac{TP}{TP+FP+FN}$, where TP stands for the number of true positives, FP for the false

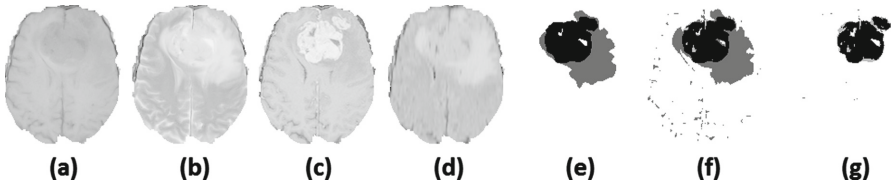


Fig. 1. Volumetric segmentation results presented in a single slice: (a)–(d) the four input channels; (e) ground truth with tumor shown in black and edema in grey; (f) and (g) results after first and second FCM stage, respectively (TP: black, FN: light grey, FP: dark grey).

Table 1. Dice scores achieved in unsupervised and various supervised settings. US stands for unsupervised processing mode, while SS1...SS12 indicate the semi-supervised processing with the number of train volumes varying between 1 and 12.

Scenario	HG01	HG02	HG03	HG04	HG07	HG11	HG14	HG15
US best	0.8925	0.7576	0.8984	0.7878	0.7476	0.7557	0.7855	0.8468
US average	0.8575	0.5318	0.8564	0.7746	0.7188	0.4480	0.7215	0.8256
SS1 average	0.8701	0.6801	0.8770	0.7786	0.7306	0.6910	0.7262	0.8359
SS2 average	0.8696	0.6872	0.8758	0.7769	0.7299	0.6956	0.7247	0.8367
SS3 average	0.8710	0.6889	0.8773	0.7773	0.7307	0.6946	0.7278	0.8371
SS4 average	0.8744	0.6912	0.8794	0.7774	0.7323	0.6974	0.7276	0.8380
SS5 average	0.8757	0.6929	0.8807	0.7775	0.7338	0.6988	0.7279	0.8382
SS6 average	0.8763	0.6940	0.8817	0.7773	0.7347	0.7002	0.7288	0.8381
SS7 average	0.8770	0.6957	0.8827	0.7769	0.7352	0.7007	0.7294	0.8382
SS8 average	0.8771	0.7001	0.8833	0.7763	0.7357	0.7006	0.7309	0.8384
SS9 average	0.8781	0.7093	0.8848	0.7758	0.7364	0.7005	0.7320	0.8385
SS10 average	0.8791	0.7207	0.8867	0.7756	0.7374	0.7012	0.7334	0.8387
SS11 average	0.8792	0.7349	0.8885	0.7754	0.7381	0.7022	0.7358	0.8391
SS12 average	0.8789	0.7491	0.8873	0.7749	0.7395	0.7016	0.7361	0.8368

positives, and FN for false negatives. Further on, the Dice score (DS) can be computed as $DS = \frac{2TP}{2TP+FP+FN} = \frac{2JI}{1+JI}$. Both indices score 1 in case of an ideal clustering, while a fully random result is indicated by a score close to zero.

3 Results and Discussion

Thirteen volumes from the BRATS 2012 data set were selected for evaluation. These underwent the proposed FCM cascade algorithm using various settings regarding fuzzy exponent and number of classes in each stage, namely $c, c' \in \{6, 7, \dots, 20\}$, and $m, m' \in \{1.5, 1.6, \dots, 2.0\}$. All these variants sum up to 8100 tests performed for each volume. The average and maximum Dice score for each volume was extracted, to be used as reference during the evaluation of the semi-supervised algorithm. Figure 1 presents a slice of a successfully segmented volume.

In order to decide the optimal number of clusters for each processing stage, we propose to employ a semi-supervised learning scenario, based on the following terms. The 13 volumes are separated into train and test data. In a given evaluation, each volume can be either train or test volume, never both. The number of train volumes, denoted by p , can vary from $p_{\min} = 1$ to $p_{\max} = 12$. At the evaluation of a certain test volume's segmentation, there are $p_{\max}!/(p!(p_{\max} - p)!)$ different possibilities to choose exactly p train volumes. Each of these cases yield a different Dice score. Average Dice scores were extracted for each test volume and each possible number of train volumes.

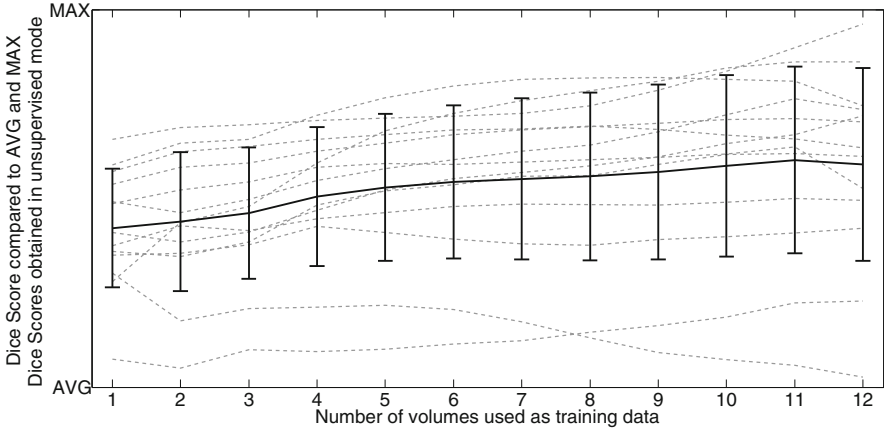


Fig. 2. Evolution of average Dice score obtained for individual test volumes, plotted against the number of train volumes (dashed lines). Average and standard deviation of the thirteen curves obtained for individual test volumes. As the number of train volumes grows, the achieved Dice score slightly tends towards the maximum.

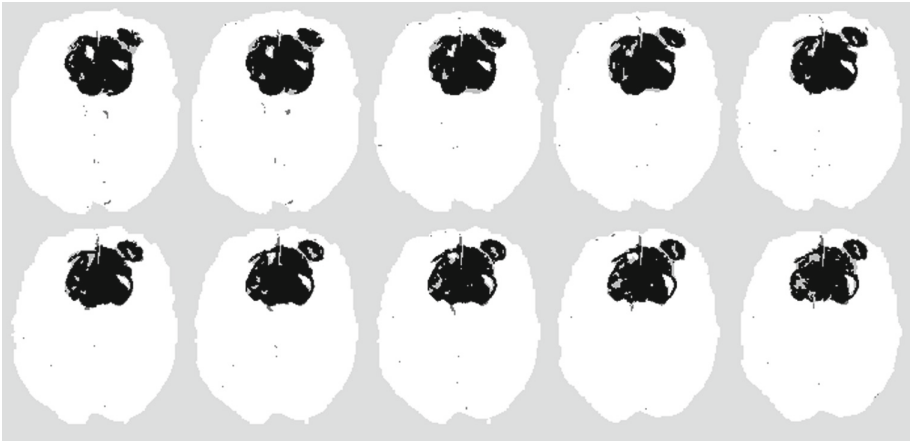


Fig. 3. Consecutive slices of a detected tumor. Black pixels are true positives, dark grey pixels represent false positives, light grey ones indicate false negatives.

Table 1 summarizes the obtained Dice scores in various circumstances. The first two rows exhibit the reference DS values (average and maximum) obtained in case of unsupervised processing. Semi-supervision induced by a given single volume, say HG01, means that we choose that parameter setting (m , m' , c , and c') which led to the best DS for volume HG01 in unsupervised processing, and applied it at the processing of the other 12 volumes. DS values shown in the SS1 row were obtained for each volume by choosing the average DS out of the 12 outcomes. Whichever is the test volume, each of the 12 others can be the inducing train volume in SS1 mode.

In semi-supervision induced by two volumes (SS2) processing mode, the parameter setting is chosen such a way that the average DS of the train volumes is greatest. This setting is applied to the other 11 volumes. Each test volume receives $12!/(10!2!) = 66$ different optimal settings from the different combinations of train volumes. Out of the 66 Dice scores were computed the values given in row SS2 of Table 1. Further rows SS3 to SS12 were obtained analogously.

The dashed curves in Fig. 2 present the evolution of relative DS achieved for each individual test volume, compared to the average (AVG) and maximum (MAX) Dice score obtained in unsupervised mode, plotted against the number of train volumes that induced the semi-supervised processing. Figure 2 also shows the global average and standard deviation of the relative Dice scores indicated by the dashed curves.

Figure 3 exhibits a tumor detected by the proposed algorithm, in consecutive slices of volume HG15. True positive pixels represented by black seem to have captured the boundary of the tumor accurately in most places. False negatives (light grey) and false positives (dark grey) are also present in a considerable proportion. So there are still challenges to improve the implemented algorithm.

As it was initially suspected, the parameters that assure best segmentation accuracy for train volumes can indeed give useful support in choosing the right parameters for the test volume. Parameters provided by the train data lead to better accuracy than the expected accuracy via randomly chosen parameters. However, not every train volume adds to the success of every test volume. Figure 2 reveals the slight improvement achieved by adding more and more volumes to the train set. Besides the initial goal of including further dozens of volumes in our study, there are several possible directions for further development of the proposed system:

1. There is a strong need to establish whether all four channels are necessary for an accurate tumor detection. If one of the channels does not really contain useful data, its presence not only hardens the computational burden, but also hinders *c*-means clustering models in providing accurate partitions.
2. Obviously not all image volumes share the same properties, not all of them are clustered best with the same parameter settings. A brief evaluation of the intensity histograms in different channels of the test data, and matching with certain previously established templates could be effective in determining, which train volumes would give best parameter setting for the test volume.
3. The current system neglects the handling of intensity non-uniformity effects. However, the effective compensation of such phenomena is usually performed simultaneously with image segmentation [13, 14].
4. The tumor detection process should be extended to the identification and separation of neighboring edema regions as well.

4 Conclusion

In this paper we proposed an automatic tumor detection and segmentation algorithm employing a fuzzy *c*-means clustering cascade algorithm and fusing

it with decision support based on prior information. The proposed algorithm was validated using 13 selected MRI volumes originating from the BRATS 2012 data set. Preliminary evaluation results revealed the ability of the algorithm to extract accurate information concerning the presence and position of the tumor. Further development directions enumerated in Sect. 3 are likely to improve the benchmarks of the algorithm.

References

1. Gordillo, N., Montseny, E., Sobrevilla, P.: State of the art survey on MRI brain tumor segmentation. *Magn. Res. Imag.* **31**, 1426–1438 (2013)
2. Asman, A.J., Landman, B.A.: Out-of-atlas labeling: a multi-atlas approach to cancer segmentation. In: 9th IEEE International Symposium on Biomedical Imaging, pp. 1236–1239. IEEE Press, New York (2012)
3. Ghanavati, S., Li, J., Liu, T., Babyn, P.S., Doda, W., Lampropoulos, G.: Automatic brain tumor detection in magnetic resonance images. In: 9th IEEE International Symposium on Biomedical Imaging, pp. 574–577. IEEE Press, New York (2012)
4. Hamamci, A., Kucuk, N., Karamam, K., Engin, K., Unal, G.: Tumor-Cut: segmentation of brain tumors on contrast enhanced MR images for radiosurgery applications. *IEEE Trans. Med. Imag.* **31**, 790–804 (2012)
5. Sahdeva, J., Kumar, V., Gupta, I., Khandelwal, N., Ahuja, C.K.: A novel content-based active contour model for brain tumor segmentation. *Magn. Res. Imag.* **30**, 694–715 (2012)
6. Njeh, I., Sallemi, L., Ayed, I.B., Chtourou, K., Lehericy, S., Galanaud, D., Hamida, A.B.: 3D multimodal MRI brain glioma tumor and edema segmentation: a graph cut distribution matching approach. *Comput. Med. Imag. Graph.* **40**, 108–119 (2015)
7. Svolos, P., Tsolaki, E., Kapsalaki, E., Theodorou, K., Fountas, K., Fezoulidis, I., Tsougos, I.: Investigating brain tumor differentiation with diffusion and perfusion metrics at 3T MRI using pattern recognition techniques. *Magn. Res. Imag.* **31**, 1567–1577 (2013)
8. Reddy, K.K., Solmaz, B., Yan, P., Avgeropoulos, N.G., Rippe, D.J., Shah, M.: Confidence guided enhancing brain tumor segmentation in multi-parametric MRI. In: 9th IEEE International Symposium on Biomedical Imaging, pp. 366–369. IEEE Press, New York (2012)
9. Yu, C.P., Ruppert, G., Collins, R., Nguyen, D., Falcao, A., Liu, Y.: 3D blob based brain tumor detection and segmentation in MR images. In: 11th IEEE International Symposium on Biomedical Imaging, pp. 1192–1197. IEEE Press, New York (2012)
10. Zhang, N., Ruan, S., Lebonvallet, S., Liao, Q., Zhou, Y.: Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation. *Comput. Vis. Image Underst.* **115**, 256–269 (2011)
11. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., et al.: The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imag.* **34**(10), 1993–2024 (2015)
12. Bezdek, J.C.: *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum, New York (1981)
13. Vovk, U., Pernuš, F., Likar, B.: A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Trans. Med. Imag.* **26**, 405–421 (2007)
14. Szilágyi, L., Szilágyi, S.M., Benyó, B.: Efficient inhomogeneity compensation using fuzzy c -means clustering models. *Comput. Meth. Prog. Biomed.* **108**, 80–89 (2012)