

The Effect of Spectral Resolution Upon the Accuracy of Brain Tumor Segmentation from Multi-Spectral MRI Data*

Ágnes Györfi, Tímea Fülöp, Levente Kovács and László Szilágyi

Abstract—Ensemble learning methods are frequently employed for brain tumor segmentation from multi-spectral MRI data. These techniques often require involving several hundreds of computed features for the characterization of the voxels, causing a rise in the necessary storage space by two order of magnitude. Processing such amounts of data also represents a serious computational burden. Under such circumstances it is useful to optimize the feature generation process. This paper proposes to establish the optimal spectral resolution of multi-spectral MRI data based feature values that allows for the best achievable brain tumor segmentation accuracy without causing unnecessary computational load and storage space waste. Experiments revealed that an 8-bit spectral resolution of the MRI-based feature data is sufficient to obtain the best possible accuracy of ensemble learning methods, while it allows for 50% reduction of the storage space required by the segmentation procedure, compared to the usually deployed featured encoding techniques.

Index Terms—Magnetic resonance imaging, spectral resolution, image segmentation, ensemble learning.

I. INTRODUCTION

The ever growing number of deployed medical imaging devices produce a continuously increasing amount of data day by day. The need for automatic data processing methods has been visibly identified. For example, the MICCAI Brain Tumor Segmentation Challenge (BraTS) that was organized yearly since 2012, led to the publication of a wide range of automated solutions for the diagnosis of brain tumors from multi-spectral MRI data [1], [2]. Recent methods proposed to the problem formulated by the BraTS challenges employ the whole arsenal of supervised and semi-supervised machine learning techniques, including ensemble learning [3], AdaBoost classifier [4], random forests [5], [6], [7],

*This project was supported in part by the Sapientia Institute for Research Programs. This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 679681). The work of T. Fülöp was supported by the Collegium Talentum 2019 Programme of Hungary. The work of L. Szilágyi was supported by the Hungarian Academy of Sciences through the János Bolyai Fellowship program.

Á. Györfi, L. Kovács and L. Szilágyi are with University Research, Innovation, and Service Center (EKIK), Óbuda University, Bécsi út 96/b, H-1034 Budapest, Hungary (phone/fax: +36-1-666-5585; e-mail: gyorfi.agnes at phd.uni-obuda.hu, {kovacs.levente, szilagyi.laszlo} at nik.uni-obuda.hu).

Á. Györfi and L. Szilágyi are also with Dept. of Electrical Engineering, Sapientia University, Calea Sighişoarei 1/C, 540485 Tirgu Mureş, Romania (phone: +40-265-206-210; fax: +40-265-206-211; e-mail: {gyorfiagnes, lalo} at ms.sapientia.ro).

T. Fülöp is with Dept. of Mathematics-Informatics, Sapientia University, Calea Sighişoarei 1/C, 540485 Tirgu Mureş, Romania (phone: +40-265-206-210; fax: +40-265-206-211; e-mail: fuloptimeal427 at gmail.com).

extremely random trees [8], support vector machines [9], [10], convolutional neural network [11], [12], deep neural networks [13], [14], [15], [16], Gaussian mixture models [17], fuzzy *c*-means clustering in semi-supervised context [18], [19], active contour models [20], [21], [22], tumor growth model [23]. For earlier solutions the reader is referred to the review paper by Gordillo *et al.* [24].

Beside creating accurate automatic processing algorithms for the image data, it is also necessary to study the efficiency of data storage. Some of the above mentioned machine learning algorithms involve several hundreds or even a couple thousands of computed features. Under such circumstances it is utmost important to use the optimal spectral resolution. Using finer resolution than the optimal one results in slower data processing. On the other hand, a too coarse resolution may damage the segmentation accuracy.

This study proposes to study the effect of color depth (also referred to as intensity resolution or spectral resolution) upon the segmentation accuracy of brain tumors from multi-spectral MRI data. The main goal is to establish, which is the optimal spectral resolution that yields the best segmentation accuracy while using the smallest number of bits to store each feature value.

The rest of this paper is structured as follows: Section II presents the data sets and data processing methods involved in the study. Section III relates on the performed experiments and their outcome. Section IV concludes this study.

II. MATERIALS AND METHODS

The problem is formulated as follows. Given an existing brain tumor detection and segmentation procedure [25] based on machine learning that uses volumetric MRI data taken from the MICCAI BraTS challenge [1]. The procedure uses an ensemble learning approach: binary decision trees (BDT) [26] are trained to separate tumor tissues from normal ones. Beside the four observed features provided by the MICCAI BraTS database, each voxel is characterized by 100 further computed features. The accuracy of the system is characterized by the average and overall Dice scores (DS), extracted according to the definitions given in Section II-D. The main goal of this study is to establish what color depth or intensity resolution is necessary to achieve the best possible performance.

A. Data

The whole set of 54 low-grade tumor volumes of the MICCAI BraTS 2016 training data set [1] was used in this study. Records contain four data channels (T1, T2, T1C,

TABLE I
INVENTORY OF COMPUTED FEATURES. OBSERVED DATA CHANNELS
WERE EQUALLY INVOLVED.

Neighborhood	AVG	MAX	MIN	MED	GRAD	GABOR	Total
$3 \times 3 \times 3$	4	4	4				12
3×3	4			4			8
5×5	4			4			8
7×7	4			4	16		24
9×9	4			4			8
11×11	4			4		32	40
Total	24	4	4	20	16	32	100

AVG - average, MAX - maximum, MIN - minimum, MED - median
GRAD - gradient, GABOR - Gabor wavelets

FLAIR) which represent the observed features for all voxels, and the annotation established by human experts. A standard automatic method was employed to register all data channels to the T1 data. Non-brain tissues were eliminated from all images. Technical details of the BraTS data are given in [1].

B. Procedure

The tumor detection and segmentation procedure employed in this study was previously described in [25], [27]. MRI data undergoes four main processing steps, which are the following:

- 1) Preprocessing, which consists of the following items: (1) intensity inhomogeneity compensation [28], [29], [30] of the MRI data accomplished with the enhanced N3 method [31]; (2) histogram normalization on all data channels via mapping to a predefined scale, accomplished with a context dependent linear transform [32] that maps the 25-percentile and 75-percentile of the intensities onto previously defined fixed values, modifies all other values accordingly, and cuts both tails of the transformed histogram at fixed thresholds.
- 2) Feature generation, which is motivated by the following: (1) neighbor voxels are likely to belong to similar tissues, so their labeling should depend on each other; (2) automatic registration is never perfect: there is no guarantee that a cubic millimeter of brain tissues is represented by the same coordinates in all observed data channels; (3) not only the observed intensities, but also the textures may contain relevant information. The additionally computed features are exhibited in Table I. Details of this feature generation process are given in our previous work [25].
- 3) Ensemble learning, which employs binary decision trees for decision making. The whole set of $n_V = 54$ LG tumor records were randomly divided into two equal subsets, which were involved in the training and testing process in turns. Ensemble consisted of a predefined number of n_T decision trees, each of which were trained to separate normal and tumor voxels within a randomly sampled set of N voxels (p_N percent negatives, $100 - p_N$ percent positives). During the training process, an entropy based criterion was deployed to select the optimal unique feature and threshold to be used by each decision node. After en-

semble training, each voxel of the test records receives a vote from each trained decision tree, and the label of the voxel is decided by the majority of the votes.

- 4) Post-processing, which improves the decision accuracy of the ensemble using a morphological criterion. The ensemble of trees provides intermediary labels only. Post-processing reevaluates the intermediary labels based on the ratio of spatial neighbors (within an $11 \times 11 \times 11$ -sized neighborhood) whose intermediary labels are positive. Post-processing has a regulatory effect upon the shape of the tumor and improves the mean values of the statistical accuracy indicators presented in Section II-D.

C. Color depth or intensity resolution

The original MRI data has each observed voxel intensity encoded in a 16-bit unsigned variable. Although this variable can store 65,536 different values, true MRI data hardly contain more than few thousand different values. Several records from the BraTS dataset only use hundreds of different intensities. During the histogram normalizing task of the preprocessing step, presented in Section II-B, the first thing is to set the target range of the normalized intensities, which is 1 to N_g , where N_g represents the number of possible grey levels. The 0 value stands for missing data or non-brain voxel within the MRI volume. We may set N_g with any value. If we intend to use a b -bit resolution, N_g should be no greater than $2^b - 1$. The predefined value of N_g defines the range of intensities for the normalized observed features, and later the range for the further 100 computed features as well.

During this study we investigated how the intensity resolution influences the accuracy of tumor segmentation within our ensemble learning framework. Choosing various values of N_g between 15 to 1023, and performing full training and testing cycles using the LG tumor volumes of the BraTS dataset, allowed us to formulate recommendations regarding the necessary number of bits to store each single value of the feature vectors.

D. Evaluation

The accuracy of the segmentation of data record i ($i = 1 \dots n_V$) is characterized by the number of true positives (TP_i), false positives (FP_i), true negatives (TN_i), and false negatives (FN_i). The Dice Score (DS) is our main accuracy indicator, which is computed for any record i as:

$$DS_i = \frac{2 \times TP_i}{2 \times TP_i + FP_i + FN_i} . \quad (1)$$

We can also extract the average DS as:

$$\overline{DS} = \frac{1}{n_V} \sum_{i=1}^{n_V} DS_i , \quad (2)$$

and the overall DS as

$$\widetilde{DS} = \frac{2 \times \sum_{i=1}^{n_V} TP_i}{\sum_{i=1}^{n_V} (2 \times TP_i + FP_i + FN_i)} . \quad (3)$$

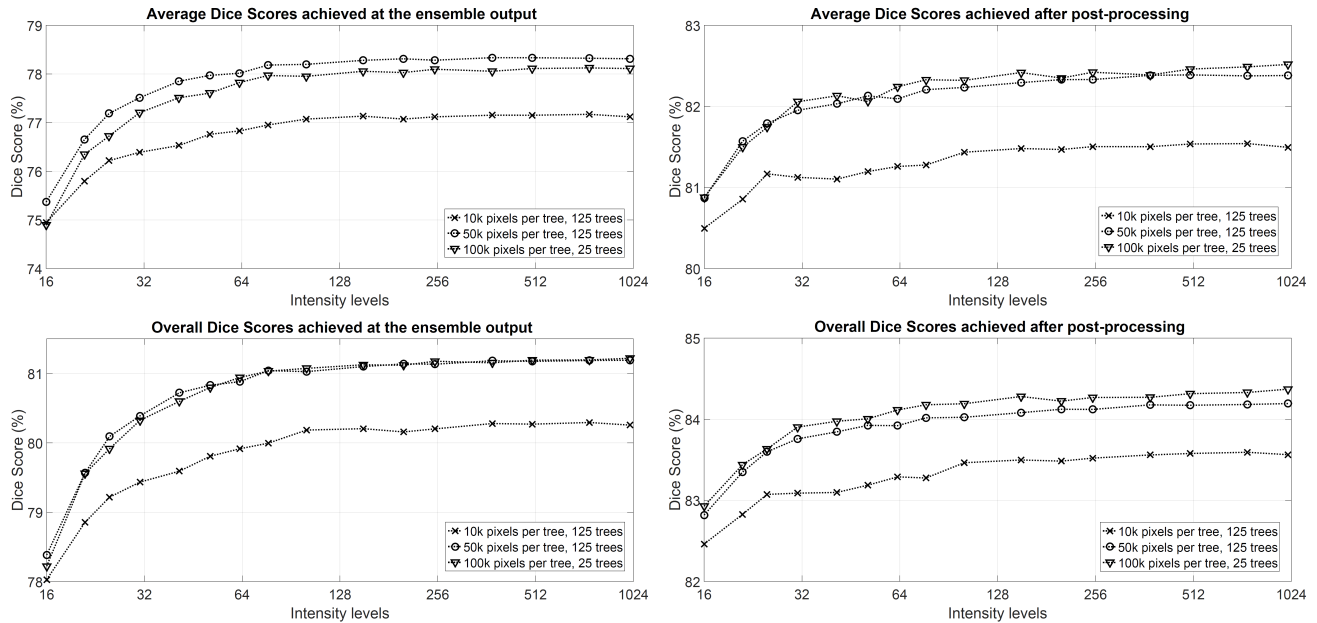


Fig. 1. Average and overall Dice scores obtained for the 54 LG tumor volumes, plotted against intensity resolution, in case of various binary decision tree ensembles.

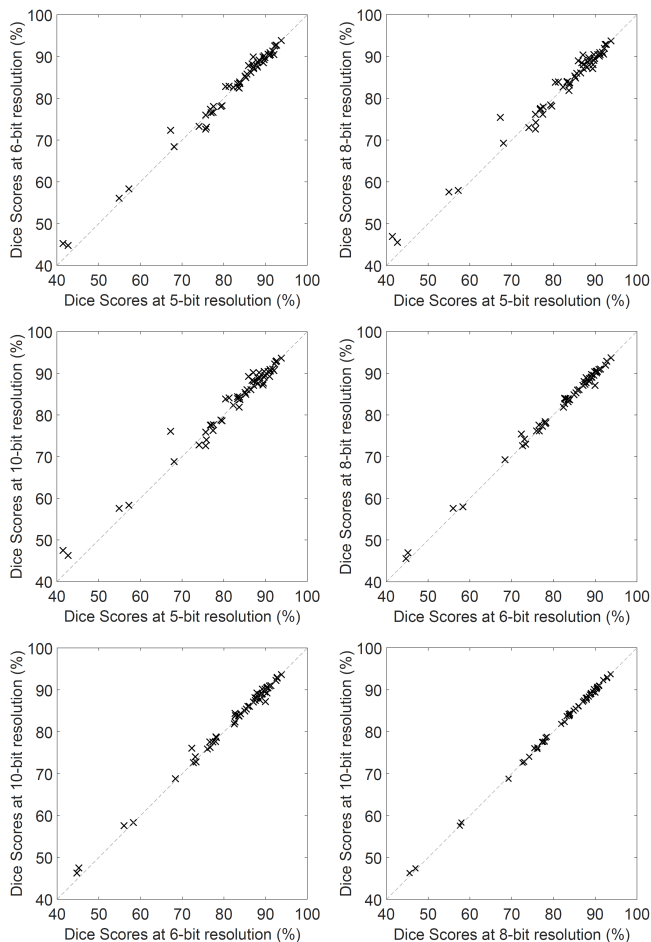


Fig. 2. Dice scores obtained for individual LG tumor volumes, plotted against each other, in all head-to-head combinations of five, six, eight and ten bits resolution.

III. RESULTS AND DISCUSSION

The 54 LG tumor records, their 4 observed and 100 computed features, were resampled at various intensity resolutions varying from 4 bits (16 intensity levels) to 10 bits (1024 intensity levels). All versions were fed to the very same ensemble learning method presented in section II, using up to 125 binary decision trees in each ensemble, each tree trained using the feature vectors of 10k to 100k voxels, out of which 93% were negatives and 7% positives.

Figure 1 presents the average and overall Dice scores obtained in various scenarios, plotted against the intensity resolution of the feature data. Dice scores seem to be stable above 8-bit resolution (256 intensity levels). However, as we decrease the color depth, the accuracy indicators drop more and more rapidly. On the other side, it is surprising that 16 intensity levels per data channel can provide acceptable Dice scores, less than 2% below the ones obtained at optimal resolution.

Figure 2 reveals us how different resolutions perform against each other, using the very same settings on the same individual LG volumes. Each panel compares two resolutions from the set of five, six, eight and ten bits, and each cross in these panels represents the Dice scores obtained for the same LG tumor volume by the two represented resolutions. Coarser resolutions usually have less accurate outcome than fine ones, but 8-bit and 10-bit resolutions perform virtually identically. Differences are better visible in case of the volumes where the Dice scores are lower.

The above results suggest that the highest achievable accuracy is guaranteed by an 8-bit resolution, so there is no need for more than a single byte to store the intensity of each voxel in the MRI volumes, after having performed the data preprocessing. Using this kind of data representation, it

is possible to save 50% of the storage space used by most MRI data formats.

IV. CONCLUSIONS

This paper investigated how much the accuracy of an ensemble learning based brain tumor segmentation procedure based on multi-spectral MRI data is influenced by the intensity resolution of the observed and computed features. The experimental study revealed that even an extremely coarse resolution of 4 bits is enough to produce acceptable results, while 8-bit resolution is enough to provide the best accuracy. This result suggests that the widely used 16-bit unsigned integer data representation is a waste of storage space, because a single byte per feature could efficiently store the full amount of relevant information.

REFERENCES

- [1] B. H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby et al., “The multimodal brain tumor image segmentation benchmark (BRATS),” *IEEE Trans. Med. Imag.*, vol. 34, pp. 1993–2024, 2015.
- [2] G. Mohan and M. Monica Subashini, “MRI based medical image analysis: Survey on brain tumor grade classification,” *Biomed. Sign. Proc. Contr.*, vol. 39, pp. 139–161, 2018.
- [3] A. Phophalia and P. Maji, “Multimodal brain tumor segmentation using ensemble of forest method,” *Proc. 3rd International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BraTS MICCAI 2017, Quebec City), Lecture Notes in Computer Science*, vol. 10670, pp. 159–168, 2018.
- [4] A. Islam, S. M. S. Reza and K. M. Iftekaruddin, “Multifractal texture estimation for detection and segmentation of brain tumors,” *IEEE Trans. Biomed. Eng.*, vol. 60, pp. 3204–3215, 2013.
- [5] N. J. Tustison, K. L. Shrinidhi, M. Wintermark, C. R. Durst, B. M. Kandel, J. C. Gee, M. C. Grossman and B. B. Avants, “Optimal symmetric multimodal templates and concatenated random forests for supervised brain tumor segmentation (simplified) with ANTSR,” *Neuroinformatics*, vol. 13, pp. 209–225, 2015.
- [6] L. Lefkovičs, Sz. Lefkovičs and L. Szilágyi, “Brain tumor segmentation with optimized random forest,” *Proc. 2nd International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BraTS MICCAI 2016, Athens), Lecture Notes in Computer Science*, vol. 10154, pp. 88–99, 2017.
- [7] Sz. Lefkovičs, L. Szilágyi and L. Lefkovičs, “Brain tumor segmentation and survival prediction using a cascade of random forests,” *Proc. 4th International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BraTS MICCAI 2018, Granada), Lecture Notes in Computer Science*, vol. 11384, pp. 334–345, 2019.
- [8] A. Pinto, S. Pereira, D. Rasteiro and C. A. Silva, “Hierarchical brain tumour segmentation using extremely randomized trees,” *Pat. Recogn.*, vol. 82, pp. 105–117, 2018.
- [9] A. Ratna Raju, P. Suresh and R. Rajeswara Rao, “Bayesian HCS-based multi-SVNN: A classification approach for brain tumor segmentation and classification using Bayesian fuzzy clustering,” *Biocyber. Biomed. Eng.*, vol. 38, no. 3, pp. 646–660, 2018.
- [10] Z. U. Rehman, S. S. Naqvi, T. M. Khan, M. A. Khan and T. Bashir, “Fully automated multi-parametric brain tumour segmentation using superpixel based classification,” *Expert Syst. Appl.*, vol. 118, pp. 598–613, 2019.
- [11] S. Pereira, A. Pinto, V. Alves and C. A. Silva, “Brain tumor segmentation using convolutional neural networks in MRI images,” *IEEE Trans. Med. Imag.*, vol. 35, pp. 1240–1251, 2016.
- [12] K. Hu, Q. H. Gan, Y. Zhang, S. H. Deng, F. Xiao, W. Huang, C. H. Cao, and X. P. Gao, “Brain tumor segmentation using multi-cascaded convolutional neural networks and conditional random field,” *IEEE Access*, vol. 7, pp. 92615–92629, 2019.
- [13] H. C. Shin, H. R. Roth, M. C. Gao, L. Lu, Z. Y. Xu, I. Noguez, J. H. Yao, D. Mollura and R. M. Summers, “Deep nonvolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning,” *IEEE Trans. Med. Imag.*, vol. 35, pp. 1285–1298, 2016.
- [14] G. Kim, “Brain tumor segmentation using deep fully convolutional neural networks,” *Proc. 3rd International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BraTS MICCAI 2017, Quebec City), Lecture Notes in Computer Science*, vol. 10670, pp. 344–357, 2018.
- [15] Y. X. Li and L. L. Shen, “Deep learning based multimodal brain tumor diagnosis,” *Proc. 3rd International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BraTS MICCAI 2017, Quebec City), Lecture Notes in Computer Science*, vol. 10670, pp. 149–158, 2018.
- [16] X. M. Zhao, Y. H. Wu, G. D. Song, Z. Y. Li, Y. Z. Zhang and Y. Fan, “A deep learning model integrating FCNNs and CRFs for brain tumor segmentation,” *Med. Image Anal.*, vol. 43, pp. 98–111, 2018.
- [17] B. H. Menze, K. van Leemput, D. Lashkari, T. Riklin-Raviv, E. Geremia, E. Alberts, et al., “A generative probabilistic model and discriminative extensions for brain lesion segmentation – with application to tumor and stroke,” *IEEE Trans. Med. Imag.*, vol. 35, pp. 933–946, 2016.
- [18] L. Szilágyi, L. Lefkovičs and B. Benyó, “Automatic brain tumor segmentation in multispectral MRI volumes using a fuzzy *c*-means cascade algorithm,” *Proc. 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2015, Zhangjiajie, China)*, pp. 285–291, 2015.
- [19] F. Özyurt, E. Sert and D. Avci, “An expert system for brain tumor detection: Fuzzy *c*-means with super resolution and convolutional neural network with extreme learning machine,” *Med Hypothesis*, vol. 134, art. no. 109433, 2020.
- [20] E. Ilunga-Mbuyamba, J. G. AvinaCervantes, A. GarciaPerez, R. J. RomeroTroncoso, H. AguirreRamos, I. CruzAceves and C. Chalopin, “Localized active contour model with background intensity compensation applied on automatic MR brain tumor segmentation,” *Neurocomput.*, vol. 220, pp. 84–97, 2016.
- [21] R. W. Ibrahim, A. M. Hasan and H. A. Jalab, “A new deformable model based on fractional Wright energy function for tumor segmentation of volumetric brain MRI scans,” *Comput. Meth. Progr. Biomed.*, vol. 163, pp. 21–28, 2018.
- [22] A. Essadique, E. Ouabida and A. Bouzid, “Brain tumor segmentation with Vander Lugt correlator based active contour,” *Comput. Meth. Progr. Biomed.*, vol. 160, pp. 103–117, 2018.
- [23] M. Lê, H. Delingette, J. Kalpathy-Cramer, E. R. Gerstner, T. Batchelor, J. Unkelbach and N. Ayache, “Personalized radiotherapy planning based on a computational tumor growth model,” *IEEE Trans. Med. Imag.*, vol. 36, pp. 815–825, 2017.
- [24] N. Gordillo, E. Montseny and P. Sobrevilla, “State of the art survey on MRI brain tumor segmentation,” *Magn. Reson. Imaging*, vol. 31, pp. 1426–1438, 2013.
- [25] L. Szilágyi, D. Iclánzan, Z. Kapás, Zs. Szabó, Á. Györfi and L. Lefkovičs, “Low and high grade glioma segmentation in multispectral brain MRI data,” *Acta Universitatis Sapientiae, Informatica*, vol. 10, no. 1, pp. 110–132, 2018.
- [26] S. B. Akers, “Binary decision diagrams,” *IEEE Transactions on Computers*, vol. C-27, pp. 509–516, 1978.
- [27] Zs. Szabó, Z. Kapás, Á. Györfi, L. Lefkovičs, S. M. Szilágyi and L. Szilágyi, “Automatic segmentation of low-grade brain tumor using a random forest classifier and Gabor features,” *Proc. 14th International Conference on Fuzzy Systems and Knowledge Discovery, Huangshan, China, 2018*, pp. 1106–1113.
- [28] U. Vovk, F. Pernuš and B. Likar, “A review of methods for correction of intensity inhomogeneity in MRI,” *IEEE Trans. Med. Imag.*, vol. 26, pp. 405–421, 2007.
- [29] L. Szilágyi, S. M. Szilágyi, B. Benyó and Z. Benyó, “Intensity inhomogeneity compensation and segmentation of MR brain images using hybrid *c*-means clustering models,” *Biomed. Sign. Proc. Contr.*, vol. 6, no. 1, pp. 3–12, 2011.
- [30] L. Szilágyi, S. M. Szilágyi, and B. Benyó, “Efficient inhomogeneity compensation using fuzzy *c*-means clustering models,” *Comput. Meth. Progr. Biomed.*, vol. 108, no. 1, pp. 80–89, 2012.
- [31] N. J. Tustison, B. B. Avants, P. A. Cook, Y. J. Zheng, A. Egan, P. A. Yushkevich and J. C. Gee, “N4ITK: improved N3 bias correction,” *IEEE Trans. Med. Imag.*, vol. 29, no. 6, pp. 1310–1320, 2010.
- [32] Á. Györfi, Z. Karetka-Mezsei, D. Iclánzan, L. Kovács and L. Szilágyi, “A study on histogram normalization for brain tumour segmentation from multispectral MR image data,” *Proc. Ibero-American Congress on Pattern Recognition (CIARP 2019, Havana), Lecture Notes in Computer Science*, vol. 11896, pp. 375–384, 2019.