Brain lesion, MRI,
classification, tensor

Konrad JACKOWSKI[1], Alexandre MANHÃES-SAVIO[1], Boguslaw CYGANEK[1,2]

AUTOMATIC SEGMENTATION OF BRAIN TUMORS USING TENSOR ANALYSIS AND MULTIMODAL MRI

Glioma detection and classification is an critical step to diagnose and select the correct treatment for the brain tumours. There has been advances in glioma research and Magnetic Resonance Imaging (MRI) is the most accurate non-invasive medical tool to localize and analyse brain cancer. The scientific global community has been organizing challenges of open data analysis to push forward automatic algorithms to tackle this task. In this paper we analyse part of such challenge data, the Multimodal Brain Tumor Image Segmentation Benchmark (BRATS), with novel algorithms using partial learning to test an active learning methodology and tensor-based image modelling methods to deal with the fusion of the multimodal MRI data into one space. A Random Forest classifier is used for pixel classification. Our results show an error rates of 0.011 up to 0.057 for intra-subject classification. These results are promising compared to other studies. We plan to extend this method to use more than 3 MRI modalities and present a full active learning approach.

1. INTRODUCTION

Gliomas are the most frequent primary brain tumors in adults. Despite advances in glioma research, patient diagnosis remains poor. The patients with high-grade glioma, the most aggressive, have a median survival rate of two years and require immediate treatment. The slower growing variants, such as low-grade astrocytomas or oligodendrogliomas, come with a life expectancy of several years so aggressive treatments are delayed as long as possible. In any case, multimodal neuroimaging is used before and after treatment to evaluate the success of the chosen treatment and the progression of the disease. This year a paper [9] presented a hierarchical majority voting method fusing several good algorithms for segmentation. Its results ranked above all individual algorithms, which indicates that methodological improvements can be achieved. Many other studies [1], [7] have presented different methods to solve this problem. In this paper we present the application of a new segmentation approach to a subset of the Multimodal Brain Tumor Image Segmentation Benchmark (BRATS) organized in conjunction with the MICCAI 2012-2015 conferences. The novelties of the methodology presented in this paper are: a) a test of an active learning approach where we only use one slice to infer the segmentation in the rest of the whole volume and; b) a tensor based data space representation

[1]ENGINE Centre, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
[2]Department of Systems and Computer Networks, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

and discriminative approach which has the potential to model the whole data space with accuracy. The active learning-based approach could allow volumetric data to be processed as a stream of slices. In the following section 2 we present the data used for this experiment, its acquisition method and the protocol for the ground truth generation. Next, in section 3 we describe the segmentation framework used to process the selected datasets. In section 4 we present the experiments that we performed and finally, we give conclusions in section 4-5.

2. MATERIALS

In this section we present the benchmark dataset used in the experiments. We describe the subjects used, a description of their clinical diagnoses, the MRI acquisition parameters and the ground truth measure protocol.

2.1. SUBJECTS

We used a random sample of 3 subjects from the BRATS 2013 benchmark.

The clinical image data consists of 65 multi-modality MR scans from glioma patients, out of which 14 have been acquired from low-grade (histological diagnosis: astrocytomas or oligoastrocytomas) and 51 from high-grade (anaplastic astrocytomas and glioblastoma multiforme tumors) glioma patients. The images represent a mix of pre- and post-therapy brain scans, with two volumes showing resections. They were acquired at four different centers: Bern University, Debrecen University, Heidelberg University, and Massachusetts General Hospital; over the course of several years, using MR scanners from different vendors and with different field strengths (1.5T and 3T) and implementations of the imaging sequences (e.g., 2D or 3D). The subject acquisitions included in this benchmark dataset all share the following four MRI modalities:

1) T1: T1-weighted, native image, sagittal or axial 2D acquisitions, with 16mm slice thickness.
2) T1c: T1-weighted, contrast-enhanced (Gadolinium) image, with 3D acquisition and 1mm isotropic voxel size for most patients.
3) T2: T2-weighted image, axial 2D acquisition, with 26mm slice thickness.
4) FLAIR: T2-weighted FLAIR image, axial, coronal, or sagittal 2D acquisitions, 26mm slice thickness.

Each subject's image volumes co-registered to the T1c MRI, which had the highest spatial resolution in most cases, and resampled all images to 1mm isotropic resolution in a standardized axial orientation with a linear interpolator. The rigid registration model used a mutual information similarity metric and three levels of resolution sampling. No attempt was made to put the individual patients in a common reference space. All images were skull stripped.

2.2. GROUND TRUTH

Here we describe the protocol for annotating the different visual structures, where present, for both low- and high-grade cases. This is further detailed in [9].

1) The edema was segmented primarily from T2 images. FLAIR was used to cross-check the extension of the edema and discriminate it against ventricles and other fluid-filled structures. The initial edema segmentation in T2 and FLAIR contained the core structures that were then relabeled in subsequent steps.

2) As an aid to the segmentation of the other three tumor sub- structures, the so-called gross tumor core including both enhancing and non-enhancing structures was first segmented by evaluating hyper-intensities in T1c (for high-grade cases) together with the inhomogenous component of the hyper-intense lesion visible in T1 and the hypo-in- tense regions visible in T1.
3) The enhancing core of the tumor was subsequently segmented by thresholding T1c intensities within the resulting gross tumor core, including the Gadolinium enhancing tumor rim and excluding the necrotic center and vessels. The appropriate intensity threshold was determined visually on a case-by-case basis.
4) The necrotic (or fluid-filled) core was defined as the tortuous, low intensity necrotic structures within the enhancing rim visible in T1c. The same label was also used for the very rare instances of hemorrhages in the BRATS data.
5) Finally, the non-enhancing (solid) core structures were defined as the remaining part of the gross tumor core, i.e., after subtraction of the enhancing core and the necrotic (or fluid-filled) core structures.

3. SEGMENTATION FRAMEWORK

In this section we are presenting the architecture of proposed system, which is subordinated to the achievement of objectives described in the introduction. Therefore, our system consists of three main modules:

1) Data acquisition module,
2) Feature extraction module, and
3) Segmentation module

3.1. DATA ACQUISITION

Original raw input material consists of series of slices from three selected MRI modalities (i.e. Flair, T1, and T2) which were described in section 2. There are two issues related with the raw data which shall be discussed.
Firstly, we are presenting a computer aided to develop a software which can help an expert (a physician) into the decision making process. Therefore, we decided design a system which allows to select only one representative slice from a series of all slices. Selection is made by a expert as the most representative one, i.e. such which allows to easily and precisely indicate tissue affected by cancer. This selected slice is further used for learning using machine learning algorithms which will subsequently automatically identify tissue state on other slices. Secondly, each image modality convey some useful data. Having an access to all of them allows to take more advantage on the potential of the MRI scanner to classify tumours. To do so, our system merges 3 of them into one compound slice.

3.2. FEATURE EXTRACTION FROM MRI SIGNALS

The processing of multidimensional signals requires proper tools. Tensor analysis, introduced and applied to signal processing and statistics, offers new insight into the multidimensional signals [8] . In this paper, tensors are applied to preprocess the fMRI signals for discovery and better discrimination of internal structures. Our approach is based on the series of developments into the multidimensional descriptions of local neighborhoods in the signal space. One of the successful approach is construction of the structural tensor, originally proposed by Bigun [2].

The method was then extended by many authors, such as in the work proposed by Luis-García et al. [6] . In this paper is further extend their approach to processing of the fMRI signals.

Our feature extraction front-end is based on computation of the extended and compact structural tensors – called EST-MRF and CST-MRF, respectively. These were used with success in many segmentation and object detection systems [4] . This type of feature extraction can be used on scalar, as well as multi-channel images. The latter option is used in our system. The intuition behind the nonlinear structural tensor is to compute averaged signal derivatives in local neighborhoods around each point. These convey information on local signal changes. More precisely, the EST-MRF is defined as follows:

$$\mathbf{E} = F_N\left(\mathbf{M}\mathbf{M}^T\right), \tag{1}$$

and

$$\mathbf{M} = \left[\begin{array}{ccccc} T_1 & T_2 & F & \bar{T}_x & \bar{T}_y \end{array}\right] \tag{2}$$

where F_N denotes the anisotropic diffusion filtering in the local neighborhood N, T_1, T_2 and F are channels of the MRI signal, while $\bar{T}_x$and $\bar{T}_y$ are average signal derivatives in the x and y directions, respectively. This way defined tensor $\mathbf{E}$ contains 15 independent components. However, in some cases these might be suppressed to more compact and sometimes even more discriminative information, called compacted structural tensor which is created from PCA processed version of the vector $\mathbf{M}$. In this case, the above equation winds up as follows.

$$\hat{\mathbf{M}} = \mathrm{PCA}\left(\mathbf{M}\right) = \left[\begin{array}{ccc} \hat{M}_1 & \ldots & \hat{M}_k \end{array}\right]^T \tag{3}$$

where k denotes a number of PCA components. This is chosen experimentally. In our experiments k was 2-3.

When (3) is inserted into (1) , then the compacted structural tensor $\mathbf{C}$ is obtained, as follows:

$$\mathbf{C} = F_N\left(\hat{\mathbf{M}}\hat{\mathbf{M}}^T\right) \tag{4}$$

For the nonlinear filter F_N in (1) the anisotropic diffusion process is employed [10]. Thanks to this, the signal is note blurred on strong signal variations, such as edges. To filter a signal T, the nonlinear heat equation is used as follows

$$\partial_t T(x,y,t) = div\left(f\left(\|\nabla T(x,y,t)\|\right)\cdot\nabla T(x,y,t)\right) \tag{5}$$

where f denotes a nonlinear control function which as its argument accepts a module of the gradient of the filtered signal T. For large gradient argument its role is to stop smoothing in this direction to avoid the smearing effect at the edge boundaries. In our system for f the Tuckey function is used. More information is provided in literature [11], [4], [5].

3.3. SEGMENTATION

The last step of data processing in our system is segmentation of slices in order to classify tumoral regions. This process is realized in two stages, i.e. training, and segmentation. As it was mentioned, one slice selected by expert is labeled as explained in subsection 2-2. We used Random Forest classifiers for training. The rest of slices are processed automatically and the pixels are marked on them by the trained classifier.

3.3.1. FEATURE NORMALIZATION

Our experience shows, that regardless of the chosen method for feature extraction (raw data from recorded slices, compacted or extended tensor), it is suggested to standardize values of attributes which form feature vector related with pixels on slices. The reason is that the structural tensor constituents vary in ranges and mixing constituents with different ranges can cause diminishing performance of classification algorithms. Therefore, we decided to normalize feature vectors.

3.3.2. PIXEL BASED CLASSIFICATION

The features extracted in the previous phase are subsequently passed to the segmentation algorithm. It aims at assigning pixels to regions which belongs to classes identified in advance by expert. As we have a multi-class classification problem we used Random Forest [3] classifier which proved its high accuracy in classification tasks.
The Random Forest classifier performs pixel based segmentation. It means, that entire slice is processed pixel by pixel and a label returned by a Random Forest is assigned to each of pixel separately. Any relationships between neighboring pixels are considered by calculating structural tensors, therefore, there is no additional analysis on relation between neighboring pixels.

4. EXPERIMENTS

We evaluated the quality of the proposed approach in two experiments carried on the benchmark dataset. We set the following objectives for our tests:

1) to compare the effectiveness of different types of data representations,
2) to examine if and how the number of classes affects the segmentation accuracy and
3) to evaluate how the segmentation accuracy varies for slices which represent images recorded in different distances from the one selected by an expert and used for learning.

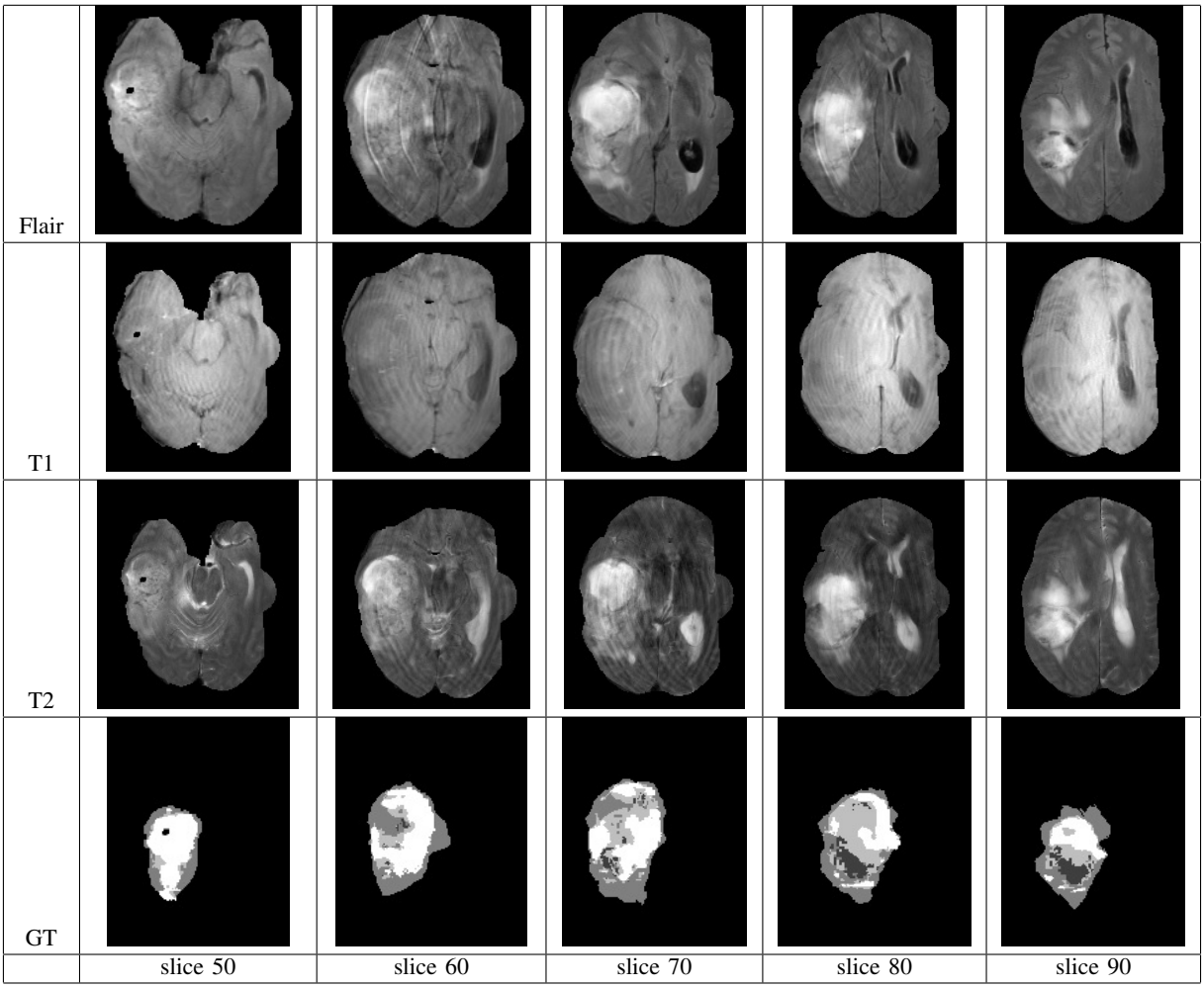
To answer the first question we compared the accuracy achieved for three methods of feature representation. Namely: Simple Stacking (i.e. stacking three types of slices), Compacted and Extended Tensor. The second question is answered by comparing the results obtained for two and five class decision tasks. Because the original data consists of 5 classes (see section 4-2), a two-class problem was generated by simple merging all cancer types into one class. Results are presented in section 4-3. The last questions are answered by comparison of accuracy for different slices in a set for the same patients. Results are presented in experiment 4-4.

4.1. EXPERIMENTAL FRAMEWORK

4.2. BENCHMARK DATA

In this experiment we randomly selected 3 patients from the BRATS 2013 dataset: 2013-pat0011, 2013-pat0013 and tcia-pat205-0001. For each of these patients we used exactly one slice for training, we chose one slice that had all the 5 classes: 90, 60 and 70, respectively. Each slice is an image of 240x240 pixels where each pixel (the third dimension) has 3 components, each corresponds to a different MRI modality: T1, T2 and Flair. Sample slices are presented in Table 1.

Table 1. Selected sample of slices recorded for patient id 0205.



4.3. EXPERIMENT 1

In these experiments we are going to test how the data representations and the number of classes affect accuracy of the segmentation. Table 2. present results, i.e. average segmentation error (number of misclassified pixels) on slices which were not used for training.

Table 2. Segmentation error for three benchmark datasets. Comparison for three method of data representations (columns), and for two and five classes decision tasks.

Dataset number of classes	Simple Stacking	Compacted Tensor	Extended Tensor
Pat_0011			
2	0.0277	0.012	0.0061
5	0.0273	0.0128	0.0072
Pat_0013			
2	0.0373	0.0312	0.0213
5	0.0369	0.0325	0.0242
Pat_0205			
2	0.0525	0.0465	0.0181
5	0.0575	0.0532	0.0361

Observations

- First of all, in all cases a simple tendency can be seen while comparing performance of three types of data representations. The worst results are always obtained by the algorithm

performing segmentation based on simple stacking raw data. Incorporating tensors allows to diminish this error significantly. It means, that analysis which takes into consideration neighborhood of the pixels allows to improve results.

- Extended tensor convey much more information than compacted one what is proved by its superior positions. Dimensionality reduction applied in compacted tensor can help while working with images of high resolution but apparently negatively affects the segmentation quality.
- Increasing number of classes caused elevating error rate only in one case (patient id 0205) and only in limited degree. It is quite positive result. We expected that there would be strict and strong correlation between number of classes and misclassification rate. Here we can see that our system is stable. Therefore we can recommend to not simplify problem. If more detailed description of the problem, i.e. decomposition of decision problem into more classes, is justified from medical point of view, it shall be done.

4.4. EXPERIMENT 2

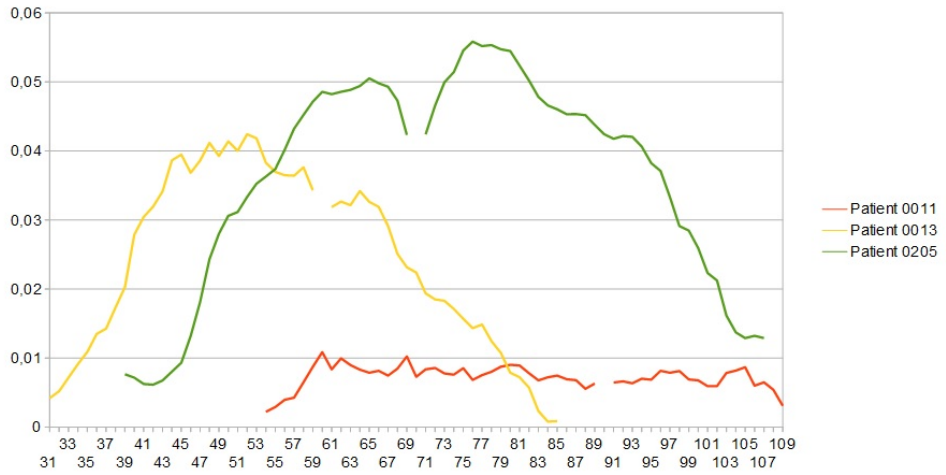


Fig. 1. Segmentation error vs number of slices. Three series represents three benchmark datasets.

Fig. 1 presents segmentation error for each of the slices used in experiments. Further explanation is needed on why the series cover different number of slices. It is so, because in the tests we selected slices which consisted of more than one class. We did so to eliminate those slices which are trivial for segmentation.
Gap in the series indicates the slice which was used for training, and therefore was eliminated from testing to avoid circularity measures.
Observations

- In two series (patient 0013, and 0205) we can see two clear tendencies.
  1) Firstly, when the distance from learning slice increases, error also increases.
  2) Next, error decreases toward zero.
- The following interpretation of aforementioned facts can be given. The pattern of cancer fluctuates from slice to slice. Therefore, in close vicinity of learning slice accuracy is high. When the distance increases, pattern changes what negatively affects the accuracy of segmentation.
- When we further increase the distance from central slice, the area occupied by the cancer gets smaller and smaller. As a result, classification process becomes easier and the error

goes to zero.

- For us, the the first observation is more important as we shall maintain stable quality of segmentation. Nonetheless, we shall emphasize that overall quality is high and their fluctuation are rather small (it never exceeds 1.5 of percent points). Some proposal how to eliminate this are provided in last section 4-5.

4.5. CONCLUSIONS

In this paper a tumor segmentation from MRI images is proposed. It is based on classification of the MRI signals preprocessed with the extended and compact versions of the structural tensor. These versions of the structural tensor allow better representation of local structures in multi-dimensional MRI signals. The classification of tensor representations is done with the random forest classifier. As shown in the experiments, the method allows fast and accurate detection of the tumor areas in the MRI brain signals. Analysis of some characteristic of the system behaviour allows to define the following direction of further works Possible directions of further works.

1) reducing fluctuation of segmentation accuracy between different slices by incorporation the following techniques: active learning methods, and machine learning algorithm which deals with concept drift and
2) improving the tensor-based modeling to deal with pixels with more than 3 components.

ACKNOWLEDGEMENT

This work was supported by EC under FP7, Coordination and Support Action, Grant Agreement Number 316097, ENGINE European Research Centre of Network Intelligence for Innovation Enhancement (http://engine.pwr.wroc.pl/)
This work was supported by the Polish National Science Centre under the grant no. DEC-2013/09/B/ST6/02264.

BIBLIOGRAPHY

[1] BAUER S., WIEST R., NOLTE L.-P., REYES M. A survey of MRI-based medical image analysis for brain tumor studies. Physics in Medicine and Biology, 2013, Vol. 58. p. R97.
[2] BIGUN J. Multidimensional orientation estimation with applications to texture analysis and optical flow. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, Vol. 13. IEEE, pp. 775–790.
[3] BREIMAN L. Random forests. Machine Learning, 2001, Vol. 45. Kluwer Academic Publishers, pp. 5–32.
[4] CYGANEK B. Object detection and recognition in digital images: Theory and practice. 2013. John Wiley & Sons, p. 552.
[5] CYGANEK B. Pattern recognition framework based on the best rank-(r 1, r 2,..., r k) tensor approximation. Computational vision and medical image processing IV: Proceedings of VipIMAGE 2013 - IV ECCOMAS thematic conference on Computational vision and medical image processing, 2013. pp. 301–306.
[6] DE LUIS-GARCIA R., DERICHE R., ROUSSON M., ALBEROLA-LOPEZ C. Tensor processing for texture and colour segmentation. Image Analysis, Lecture Notes in Computer Science, Vol 3540, 2005. pp. 1117–1127.
[7] GORDILLO N., MONTSENY E., SOBREVILLA P. State of the art survey on MRI brain tumor segmentation. Magnetic Resonance Imaging, Oct. 2013, Vol. 31. pp. 1426–1438.
[8] LATHAUWER D. L. Signal processing based on multilinear algebra. 1997. PhD dis-sertation, Katholieke Universiteit Leuven.
[9] MENZE B., JAKAB A., BAUER S. E. A. The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Transactions on Medical Imaging, Oct. 2015, Vol. 34. pp. 1993–2024.
[10] PERONA, P. M. J. Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, Jul 1990, Vol. 12. IEEE, pp. 629–639.
[11] SAPIRO G. Geometric partial differential equations and image analysis. 2006. Cambridge University Press, p. 385.