Graph-based Detection, Segmentation & Characterization of Brain Tumors

Sarah Parisot, Hugues Duffau, Stéphane Chemouny, Nikos Paragios

Graph-based Detection, Segmentation & Characterization of Brain Tumors

Sarah Parisot[1,2,4] *, Hugues Duffau[5], Stéphane Chemouny[4], Nikos Paragios[1,2,3]

[1] Center for Visual Computing, Ecole Centrale de Paris, France
[2] Equipe Galen, INRIA Saclay, Ile-de-France, France
sarah.parisot@ecp.fr
[3] Université Paris-Est, LIGM (UMR CNRS), Center for Visual Computing, Ecole des Ponts ParisTech, France
nikos.paragios@ecp.fr
[4] Intrasense SAS, Montpellier, France
chemouny@intrasense.fr
[5] Département de Neurochirurgie, Hopital Gui de Chauliac, CHU Montpellier, France
h-duffau@chu-montpellier.fr

Abstract

In this paper we propose a novel approach for detection, segmentation and characterization of brain tumors. Our method exploits prior knowledge in the form of a sparse graph representing the expected spatial positions of tumor classes. Such information is coupled with image-based classification techniques along with spatial smoothness constraints towards producing a reliable detection map within a unified graphical model formulation. Towards optimal use of prior knowledge, a two layer interconnected graph is considered with one layer corresponding to the low-grade glioma type (characterization) and the second layer to voxel-based decisions of tumor presence. Efficient linear programming both in terms of performance as well as in terms of computational load is considered to recover the lowest potential of the objective function. The outcome of the method refers to both tumor segmentation as well as their characterization. Promising results on substantial data sets demonstrate the extreme potentials of our method.

1. Introduction

Tumor detection, and in particular at early stage is of extreme clinical interest. Recent development of imaging as well as contrast-enhanced modalities have made possible the in-vivo/non-invasive detection and characterization of tumors. This information is critical to physicians towards intervention/therapy planning as well as for evaluating different therapeutic strategies. However, the problem is ill-posed due to the enormous variability of tumors both in terms of location as well as in terms of geometric characteristics and progression. Contrast enhanced imaging alleviates the problem to certain extend, but still introduces significant appearance variability. Conventional medical image segmentation techniques adopt smoothness constraints and prior knowledge to overcome the ill-posedeness of the task, however this is far from being trivial in the case of tumors presence modeling.

Prior art in tumor brain detection is limited. In [15], the lesions are detected as outliers with respect to the normal tissue brain characteristics and a healthy brain atlas is used for spatial and geometric constraints. This kind of method assumes that the lesions have significant intensities differences with the rest of the image. In [11] a probabilistic brain atlas is modified to include prior probabilities for tumor (enhancing areas) & edema (fraction of the white matter prior probability). This method fails in case of large deformations and requires multiple modalities. [6] alternates between statistical classification of different tissue types and registration of the data with a manually segmented anatomical atlas. This assumes strong homogeneity in the tumor appearance. Another approach [1] performs non rigid registration while simulating a manually seeded tumor growth.

The main limitation of these methods is that prior knowledge is encoded in a rather global manner, which is problematic for two reasons. First, given the diversity of brain tumors, the number of samples needed for their statistical characterization is important. On top of that, global models are not adequate for tumor modeling since usually they have

*This work was supported by ANRT (grant 147/2010), Intrasense and the European Research Council Starting Grant Diocles (ERC-STG-259112).

a rather systematic presence locally [3, 9] - where eventually it might be feasible to create meaningful prior models - but this is not the case globally.

This paper proposes a novel prior representation for tumor detection and segmentation. This is achieved by seeking a sparse graph where tumors have been clustered according to their spatial proximity [13]. Each class corresponds to a distinct tumor spatial behavior and is determined through unsupervised clustering. This model is propagated to a Markov Random Field. The data support is encoded through modern machine learning techniques (boosting), while the prior term aims to determine the type of tumor and its optimal spatial position. The unknown variables of the model refer to the binary segmentation map (tumors versus healthy tissue) and the associated tumor characterization. We evaluate the performance of the method in the context of low-grade gliomas.

The reminder of this paper is organized as follows. In section 2, we present the prior model. Section 3 is dedicated to the segmentation model while experimental results and validation are part of Section 4. Discussion concludes the paper.

2. Tumor Characterization & Representation Using Sparse Graphs

Statistical modeling of tumors presence can be achieved with any of the conventional/advanced dimensionality reduction techniques (PCA, ICA, IsoMap). [13] demonstrated that there exists preferential locations for low-grade gliomas in the brain. This kind of atlas gives useful information on where the tumors are likely to appear and can be a powerful tool for tumor segmentation. In this paper, we construct a similar atlas following [13]'s methodology and use it as a spatial position prior for tumor segmentation. Let us briefly review the material presented in [13] towards a self-contained presentation of the prior.

Let us consider a set $S_i, i \in [1..n]$ of segmentation maps, obtained through manual segmentation of n MRI FLAIR volumes of n different patients. In order to compensate the inter-patient anatomical variability and to be able to compare the tumors' positions, we first perform affine registration of all volumes (and segmentation maps) to the same tumor free reference pose.

The first step for the construction of the atlas is to evaluate the proximity between tumors. This is done using the Mahalanobis distance[14]:

$$\begin{aligned} d_M(S_i, S_j) &= \sqrt{(\bar{\mathbf{x}_\mathbf{i}} - \bar{\mathbf{x}_\mathbf{j}})^T \Sigma^{-1} (\bar{\mathbf{x}_\mathbf{i}} - \bar{\mathbf{x}_\mathbf{j}})} \\ \text{where} \qquad \Sigma &= \frac{(n_i - 1)\Sigma_i + (n_j - 1)\Sigma_j}{n_i + n_j - 2} \end{aligned} \tag{1}$$

where $\bar{\mathbf{x}_\mathbf{i}}$ and $\bar{\mathbf{x}_\mathbf{j}}$ are the center of mass of S_i and S_j, Σ_i and Σ_j are the covariance matrices of voxels coordinates of S_i and S_i while n_i and n_j are the number of voxels in tumors S_i and S_j respectively. Basically, it measures the amount of overlap between the tumors that are approximated as ellipsoids, taking into account the sizes, positions and orientations of the tumors in the brain. This distance is well adapted to the kind of edema free tumor we work on, and could easily be changed in case of a different pathology. The Mahalanobis distance is computed for all pairs of tumors in our data-set, resulting in a similarity matrix. This is used to construct a complete graph where each node represents a tumor (i.e. a patient) and the arcs' strength corresponds to the Mahalanobis distance value.

In order to identify the preferential locations, we seek to regroup nearby tumors whose positions can statistically be expressed by a single node, using a recent clustering algorithm [7] where the number of clusters k, clusters centers $\bar{c_1}, \ldots, \bar{c_k}$ and remaining nodes assignments $c_1, \ldots, c_k$ are to be determined. This is a minimization problem:

$$\min_k \min_{\bar{c_i}} \min_{l_j} \left(\sum_{i=1}^{k} V_g(\bar{c_i}) + \alpha \sum_{i=1}^{k} \sum_{j=1}^{n} \delta(c_i - l_j) d(\bar{c_i}, S_j) \right) \tag{2}$$

where l_j is the assignment of observation j and α is a constant coefficient balancing the contributions of the two terms. The second term in the equation simply assigns a node that hasn't been selected as a center to the closest cluster. In order to avoid the trivial solution of designating each node as a cluster center, a penalty term V_g is introduced.

$$V_g(S_i) = \sum_{j=1}^{n} d(S_i, S_j) \tag{3}$$

This term aims at selecting as centers the nodes that have strong overlaps with the rest of the nodes.

This algorithm relies on reformulating (2) as an equivalent integer program of the following form:

$$\text{PRIMAL-IP} \equiv \min_{\mathbf{x}} \sum_i V_g(S_i) x_i + \sum_{j,i} d(S_j, S_i) x_{ji} \tag{4}$$

$$\text{s.t.} \sum_i x_{ji} = 1 \tag{5}$$

$$x_{ji} \leq x_i \tag{6}$$

$$x_{ji}, x_i \in \{0, 1\} \tag{7}$$

In the above formulation each binary variable x_{ji} indicates whether observation j has been assigned to node i or not, while binary variable x_i indicates whether node i has been chosen as a central node or not. Such a minimization problem is solved through a one-shot optimization using LP-programming and the notion of stability is driving the selection of cluster centers.

The clustering results will heavily depend on the value of α. 3 cluster validity indices are used to evaluate the value

of α that yields the best clustering. As suggested in [13] we adopt well known indices to determine the optimal clustering. The Dunn index [4] aims at identifying compact and well separated clusters by comparing the biggest distance intra-cluster to the smallest distance inter-clusters.

$$D = \min_{i\in[1:k]}\{\min_{j\in[1:k],j\neq i}\{\frac{d(\bar{c_i},\bar{c_j})}{d_{max}}\}\} \quad (8)$$

where k is the number of clusters, d_{max} is to the maximal distance of a node to the center of the cluster it belongs to (distance intra-cluster), and $d(\bar{c_i},\bar{c_j})$ is the Mahalanobis distance between the centers $\bar{c_i}$ and $\bar{c_j}$ of clusters i and j (distance inter-clusters). The best clustering will correspond to a Dunn index that is maximum (small distance intra-cluster and high distance inter-clusters).

The Davies-Bouldin index [2], aims at identifying compact and well separated clusters. A measure of similarity between clusters is defined

$$R_{i,j} = \frac{\sigma_i + \sigma_j}{d(\bar{c_i},\bar{c_j})} \quad (9)$$

The Davies-Bouldin index computes the maximum similarity:

$$DB = \frac{1}{k}\sum_{i=1}^{K}\max_{j\in[1:k],j\neq i} R_{i,j} \quad (10)$$

where σ_i is the average distance between all samples in cluster c_i and the center $\bar{c_i}$ of the cluster. The best clustering corresponds to a Davies-Bouldin index that is minimum (little similarities between clusters).

Eventually, the Silhouette index [16] evaluates how well each sample in the data-set fits in its assigned cluster.

$$s(S_i) = \frac{b(S_i) - a(S_i))}{\max(a(S_i), b(S_i))} \quad (11)$$

where $a(S_i)$ is the average distance between sample S_i and all the remaining samples in the S_i's cluster c_i. $b(S_i)$ is the minimum average distance between S_i and all of the elements clustered in $C_j, (j = 1, ..., k; j \neq i)$. $s(S_i)$ values vary between -1 and 1. A value close to 1 means the cluster assignment is adequate, close to 0 suggests that the sample is equally far away from 2 clusters, while a value close to -1 imply misclassification. To evaluate the quality of the clustering, we compute the global silhouette index:

$$GS = \frac{1}{K}\sum_{j=1}^{K}\frac{1}{n_j}\sum_{i=1}^{n_j} s(S_i) \quad (12)$$

where n_j is the number of elements in cluster c_j. The best clustering will correspond to the maximum global silhouette index.

The optimal network representation will be selected according to the optimal cluster validity indices.

3. Tumor Characterization & Detection

Let us consider without loss of generality that the outcome of the sparse graph representation consists of m clusters, and $c_1, \cdots, c_m$ being the labels corresponding to these clusters. Let us consider for each cluster that a statistical model has been build with respect to the tumor presence at a given voxel denoted with $p_{c_i}(\mathbf{x})$. This model can simply be constructed by using the empirical distribution of tumor detections per voxel withing the cluster. Let S_{c_i} be the representative tumor (cluster center) corresponding to the cluster c_i with a binary label associated to it. Last but not least let us consider that a classifier $f(\mathbf{x})$ has been built acting on features derived from the image space, separating healthy versus tumor voxels. We denote $p_{tm}(f(\mathbf{x})), p_{bg}(f(\mathbf{x}))$ the probabilities for voxel $\mathbf{x}$ of belonging to the tumor and background class respectively. Without loss of generality we assume that a common classifier has been built for all tumors but individual regressors per class could be considered as well.

The problem of detection, characterization & segmentation in a new image can be expressed using two random variables $(\theta(\mathbf{x}), \omega(\mathbf{x})) \in \{c_1, \cdots, c_n\}\mathrm{x}\{bg, tm\}$ defined at the voxel level. The first label acts on the entire volume and characterizes the type of tumor (i.e. assigns the tumor to a cluster), while the second is acting on the voxel level and makes a binary call depending on the presence of tumor or not. Clinically, such an assumption is valid since it is rather unusual to observe tumors of different types in the same subject. Therefore, we seek to assign a label $\mathbf{l}(\mathbf{x}) = \{\theta(\mathbf{x}), \omega(\mathbf{x})\}$ for each voxel of the volume. We reformulate this labeling problem as a Markov Random Field on $\mathbf{l}$ where the 2 graphs are interconnected:

$$E(\mathbf{l}) = \sum_{\mathbf{x}} V_{sing}(\mathbf{l}(\mathbf{x})) + \sum_{\mathbf{x}}\sum_{\mathbf{y}} V_{pair}(\mathbf{l}(\mathbf{x}), \mathbf{l}(\mathbf{y})) \quad (13)$$

Let us now proceed with the definition of the singleton term. It consists of 3 different potentials acting on the segmentation and characterization spaces:

$$V_{sing}(\mathbf{l}(\mathbf{x})) = V_{sg}(\omega(\mathbf{x})) + \lambda V_{ch}(\theta(\mathbf{x})) + \mu V_{ch,sg}(\theta(\mathbf{x}), \omega(\mathbf{x})) \quad (14)$$

where λ and μ are 2 constant parameters determining the relative importance of the potentials.

The first term, V_{sg}, acts on the segmentation space $\omega(\mathbf{x})$. It makes use of the classifier's output to label a voxel as tumor or background: voxels with a high classification score will most likely be labeled tumor.

$$V_{sg}(\omega(\mathbf{x})) = -log(p_{\omega(\mathbf{x})}(f(\mathbf{x}))) \quad (15)$$

In order to determine the tumor position in the absence of complete segmentation, let us consider the Heaviside dis-

tribution:

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

Classification responses with strong tumor support provide valuable information on the type of the tumor. Ideally, it is expected that such responses should be in agreement with the segmentation map representing the cluster center of the associated tumor type. Based on the above, the tumor type support is defined as:

$$V_{ch}(\theta(\mathbf{x})) = 1 - \delta(H(f(\mathbf{x}) - T_{det}), S_{\Theta(\mathbf{x})}(\mathbf{x})) \quad (17)$$

with T_{det} being a value experimentally determined from the training set. The second singleton potential V_{ch} acts on the tumor characterization space $\theta(\mathbf{x})$. This term aims at imposing the tumor type which optimally overlaps with the strong classification decisions and once considered in the global formulation; can be seen as an empirical estimation of the hamming distance between the representative tumor of the cluster and the one detected in the new image.

Both those terms will enable to label the voxels individually. In the 2 cases, we include pairwise costs in order to add neighborhood information:

$$V_{pair}(\mathbf{l}(\mathbf{x}), \mathbf{l}(\mathbf{y})) = V_{ch,ch}(\theta(\mathbf{x}), \theta(\mathbf{y})) + V_{sg,sg}(\omega(\mathbf{x}), \omega(\mathbf{y})) \quad (18)$$

In the segmentation space, we want to avoid isolated detections as well as to impose local consistency of the segmentation. To this end, we include a pairwise term on $\omega(\mathbf{x})$ adopting the conventional Potts model, or:

$$V_{sg,sg}(\omega(\mathbf{x}), \omega(\mathbf{y})) = \begin{cases} 0, & \text{if } \omega(\mathbf{x}) = \omega(\mathbf{y}) \\ \beta, & \text{otherwise} \end{cases} \quad (19)$$

While we seek a binary output on the segmentation space, we aim at labeling the entire volume with the same $\theta(\mathbf{x})$ value. Indeed, we want to assign the whole image to the same cluster of tumors. This is imposed by a second pairwise term acting on the characterization space that forces the same labeling on the whole image:

$$V_{ch,ch}(\theta(\mathbf{x}), \theta(\mathbf{y})) = \begin{cases} 0, & \text{if } \theta(\mathbf{x}) = \theta(\mathbf{y}) \\ \infty, & \text{otherwise} \end{cases} \quad (20)$$

Last, let us introduce the singleton term that actually acts as a prior and couples the two graphs:

$$V_{ch,sg}(\theta(\mathbf{x}), \omega(\mathbf{x})) = \begin{cases} p_{\theta(\mathbf{x})}(\mathbf{x}), & \text{and } \omega(\mathbf{x}) = bg \\ 1 - p_{\theta(\mathbf{x})}(\mathbf{x}), & \text{and } \omega(\mathbf{x}) = tm \end{cases} \quad (21)$$

A probability map is constructed for each cluster, describing the distribution of tumor appearances per voxel. The characterization term V_{ch} enables to identify which probability map is to be used. Then, the prior term will either compete with the classifier's information (term V_{sg}) in case of detections that do not correspond to frequent tumor appearances, or support the classification decisions if they correspond to positions where a lot of tumors have appeared. For instance, false positives that are not in the vicinity of the tumor are likely to be eliminated. While we could have used $-log(p_{\theta(\mathbf{x})})$, it would penalize tumors that do not fit completely in their assigned cluster probability distribution.

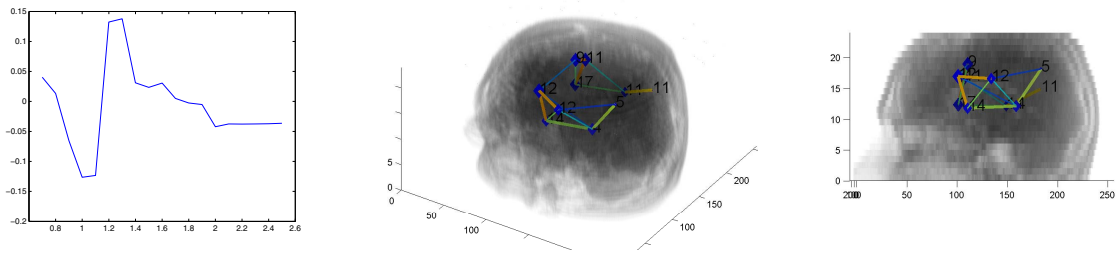
The resulting formulation can be optimized using conventional discrete optimization techniques. We adopt the FastPD [8] algorithm due to the fact that it offers the best compromise between computational complexity and performance.

4. Experimental Validation

4.1. Data-set and preprocessing

Our data set consisted of 113 3D MRI FLAIR images of 113 different patients with low grade gliomas. We work solely with 3D MRI FLAIR images of low-grade glioma, since this modality offers the best contrast between healthy tissue and low-grade gliomas' tumorous tissue. The patient age ranged from 21 to 65 years, and tumor size between 3.5 and 250 cm^3. Each image had been manually annotated by experts to indicate the position of the tumor. The image size varied from 256x256x24 to 512x512x33. The voxel resolution ranged from 0.4x0.4 to 0.9x0.9 mm^2 in the (x,y) plane and 5.3 to 6.4 mm in the z plane. The most frequent size and resolution were 256x256x24 and 0.9x0.9x5.5 mm^3. The reference pose for registration was a FLAIR image (size 256x256x24, resolution 0.9x0.9x5.5 mm^3) of a tumor free brain. Since spatial position of the voxels is a key element in our work, we perform affine registration on all our data-set (images and segmentation maps) to the atlas. One can note that it is easy to study the segmentation results directly on the patient's space by applying the inverse transformation.

One of the biggest issues involving MRI images is that the intensity can considerably vary from one image to another, and even within one image. Many intensity regularization algorithms have been proposed [12], but they are more adapted to healthy brains. A complex algorithm could diminish the contrast between tumor and brain tissue by assuming the contrast enhancement corresponding to the tumor is due to MRI inhomogeneity. As a result, tumors with low contrast enhancement could no longer be detectable. We adopt a simple regularization method based on image statistics. Our goal is to have images intensities in the same range. To this end, we use the median intensity μ and interquartile range IQR without taking into account background voxels of the image (that would influence the median value). We set the same median and interquartile range



(a) (b) (c)

Figure 1: (a) Mean value of the combination of the 3 criteria in function of $alpha$ over the 107 graphs. (b,c) Example clustered network representation

to all images by computing the new intensity as

$$I_{new} = \frac{I - \mu}{IQR} \tag{22}$$

Among our 113 images, 40 were used to train the classifier, and segmentation results were evaluated on the 73 remaining images.

4.2. Prior Construction

A high number of images as well as a limited tumor size were necessary to obtain a statistically significant graph structure. The prior was constructed using leave-one-out cross validation, so that the maximum number of images is used for prior construction. 6 Images whose tumor size was larger than 125 cm^3 were excluded as they could alter the structure of the graph.

For each sample, the graph was constructed using the 106 remaining images. In order to select the best clustering, we have clustered the graphs and computed the cluster validity indices for several α values. The mean value of the indices was computed in order to select the α value that represent well the data. As we can see in [Fig. (1a)], all the considered criteria have produced a remarkable optimum in terms of α, which corresponds to a set of 10 clusters. An example clustered graph is shown in [Fig. (1b)] and [Fig. (1c)]. For each cluster c_i and sample, we constructed a probability map describing the frequency of tumor appearances per voxel as:

$$p_{c_i}(\mathbf{x}) = \frac{1}{n_i} \sum_{j=1}^{n_i} S_i^{c_i}(\mathbf{x}) \tag{23}$$

where $S_i^{c_i}$ are the binary maps of the tumors clusters in c_i (value 1 for tumor and 0 for background) and n_i is the number of elements in cluster c_i. The prior used for the 6 excluded images is randomly selected among the 107 graphs.

4.3. Classification Likelihoods

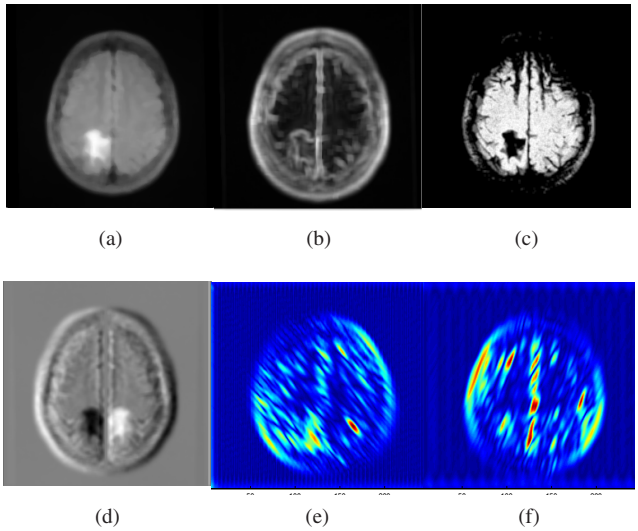
We will rely on boosting to build a tumor versus background classifier $f(\mathbf{x})$. It is based on the idea that a combination of weak classifiers such as decision stumps can create a strong classifier. We use the Gentle Adaboost algorithm [5] and 40 randomly selected images (used for prior construction) to train the classifier with the following features:

- **Intensity based features:** Intensity enhancement is the main difference from tumor to normal brain tissue. Low grade gliomas appear brighter than brain tissue on FLAIR images. However, intensity alone is not sufficient, as there exist overlapping intensity values with normal brain tissue and variable intensities within the tumor. We use 9*9*5 intensity patches to add information about the neighborhood of the voxel. We also computed the median, standard deviation and entropy of intensity patches of size k*k*3, where k=[3,5,7]. Examples of intensity based features are shown in Fig. 2a, 2b and 2c

- **Gabor features:** Gabor filters [10] have been commonly used for texture segmentation. Tumor and normal brain tissue have a very similar texture, therefore Gabor features cannot be used for detecting the tumor. We use them on 2 scales and 10 orientations mostly in order to gain information about the main structure of the brain. Gabor features for different scales and orientations are shown in Fig. 2e and Fig. 2f

- **Symmetry feature:** One interesting characteristic of the brain is its approximate symmetry. This is an important asset for tumor segmentation since a tumor will introduce a notion of dissymmetry in the brain. Our images being affinely registered to the reference pose, their symmetry plane is roughly equivalent to the reference pose's. Let Π be the reference pose's symmetry plane and voxel p_Π the symmetric to voxel p with respect to Π. We estimate a symmetry measure:

$$S(p) = \frac{1}{N} \sum_{N(p)} I(p) - \frac{1}{N} \sum_{N(p)} I(p_\Pi) \tag{24}$$



(a) (b) (c)

(d) (e) (f)

Figure 2: Examples of features used for boosting training: (a) median filter, (b) standard deviation, (c) entropy, (d) symmetry, (e,f) Gabor features.

where N(p) is a neighborhood patch of p which role is to compensate the approximate symmetry plane, and N is the number of voxels in N(p). [Fig. 2] shows examples of features. An example of symmetry feature is shown in Fig. 2d

At each iteration, the boosting algorithm selects a feature $F(\mathbf{x}, i)$ and a threshold tsh and builds a weak classifier $f_i(\mathbf{x})$:

$$f_i(\mathbf{x}) = \begin{cases} = 1, & \text{if } F(\mathbf{x}, i) \leq tsh \\ = -1, & \text{otherwise} \end{cases} \quad (25)$$

The strong classifier will be a linear combination of the weak classifiers. As a result, we get for each voxel, a score $f(\mathbf{x})$ of confidence of being tumor and convert it to a probability:

$$p_{tm}(f(\mathbf{x})) = \frac{1}{1 + exp(-2f(\mathbf{x}))} \quad (26)$$

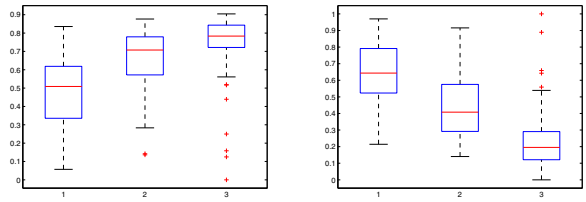
We also compute the background probability as:

$$p_{bg}(f(\mathbf{x})) = 1 - p_{tm}(f(\mathbf{x})) \quad (27)$$

4.4. Segmentation results

To evaluate the quality of the segmentation, we compare the automatic segmentation A to a manual segmentation M made by a low-grade gliomas expert using the Dice value and the rate of false positives:

$$D = \frac{2\|M \cap A\|}{\|A\| + \|M\|} \qquad FP = \frac{\|A\| - \|M \cap A\|}{\|A\|} \quad (28)$$



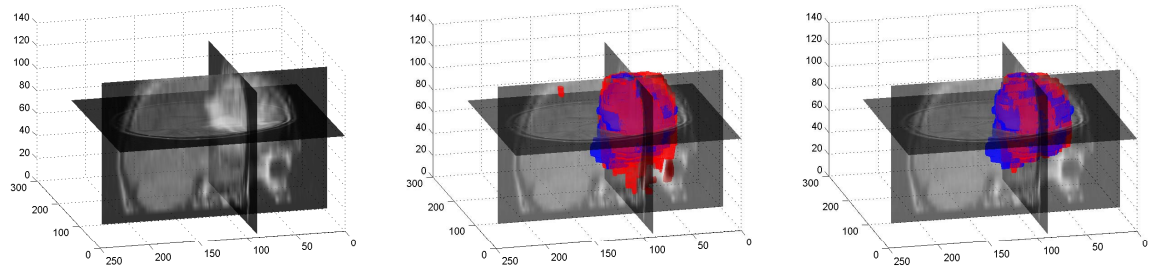
(a) (b)

Figure 3: Boxplots of the Dice values (a) and rate of false positives. From left to right: segmentation results with boosting only, boosting and pairwise regularization, boosting, prior and pairwise regularization.

We test the boosting results and optimize the prior's parameters on the images that belong to the clustered graph and were not used for boosting learning. Using our framework, all images were automatically assigned to the right cluster. We noticed experimentally that the singleton potential associated with cluster assignment (V_{dt}) had to be given a much higher importance than the other singletons in order to achieve correct cluster assignment. Therefore, we set the λ parameter to 300, while the high boosting score threshold T was set to 1.5. As for the prior term, we observed the best results by setting $\mu = 2$. Eventually, pairwise cost β was set to 1.

We then test our framework on the 73 images that were not used for boosting learning using the optimized parameters. The mean Dice increases from 48% (boosting only), 65% with pairwise regularization to 74%. As for the false positive rates, boosting associated with pairwise regularization alone yielded a false positive rate of 45%. It dropped to 24% when adding the prior. [Fig. (3)] shows boxplots of the dice values and false positive rate for both training and test set with and without prior while [Fig. (5)] shows examples of different stages of segmentation on 3D volumes slices, and [Fig. 4] shows a 3D representation of a segmentation. No comparisons with global statistical modeling methods are reported since this case was addressed in [13].

The proposed method led to very promising experimental results in challenging data sets. Successful segmentations include tumors with some necrosis, fuzzy boundaries, low contrast enhancement and different sizes. We could observe however that the prior didn't perform well in the case of a very big tumor ($> 200cm^3$) . This is due to the fact that the tumor is too big to be covered by a single cluster and could be compensated by combining the information from a couple clusters. The tumor identification (cluster assignment) could also give interesting insight on the future development of the tumor, as low-grade gliomas could have a location dependent behavior.



(a) (b) (c)

Figure 4: 3D representation of an example segmentation. (a) Original image, (b) Segmentation with boosting and pairwise regularization, (c) segmentation with prior. Both segmentations (red) are superimposed to the manual segmentation (blue)

5. Discussion

In this paper we have proposed a novel way to encode prior knowledge in tumor segmentation, making use of the fact that the tumors tend to appear in the brain in preferential locations. We combine an image-based detections schema with identification of the tumor's corresponding preferential location, which is associated to a specific spatial behaviour.

Future work involves better modeling of the prior knowledge through a more appropriate geometric modeling of tumor proximity that better encodes registration errors. The segmentation component of the method could greatly benefit from higher order interactions between voxels. The idea is to express cluster geometric behavior as a collection of higher order cliques and then impose on the segmentation process consistency for all voxels belonging to these cliques. Such an approach might also inherit pose invariance and eliminate the need of the affine registration step. The use of the proposed framework for clinical purposes is in progress in particular for tumor progression characterization. In terms of computer vision problems, the proposed formulation is well suited for part-based detection, segmentation and characterization in particular when referring to structures with parts underlying inconsistent motion. Body pose estimation is an example where such method could be used towards combined segmentation and action recognition.

References

[1] M. Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J. Villemure, and J.-P. Thiran. Atlas-based segmentation of pathological mr brain images using a model of lesion growth. *IEEE Transactions on Medical Imaging*, 23(10):1301–1314, 2004.

[2] D. L. Davies and D. W. Bouldin. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-1:224–227, 1979.

[3] H. Duffau and L. Capelle. Preferential brain locations of low-grade gliomas. *Cancer*, 100:2622–2626, Jun 2004.

[4] J. C. Dunn. Well separated clusters and optimal fuzzy-partitions. *Journal of Cybernetics*, 4:95–104, 1974.

[5] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: a statistical view of boosting. *The Annals of Statistics*, 38(2):337?407, 2000.

[6] M. Kaus, S. Warfield, A. Nabavi, P. Black, F. Jolesz, and R. Kikinis. Automated segmentation of mr images of brain tumors. *Radiology*, 218(2):586–591, 02 2001.

[7] N. Komodakis, N. Paragios, and G. Tziritas. Clustering via lp-based stabilities. In *NIPS*, pages 865–872, 2008.

[8] N. Komodakis, G. Tziritas, and N. Paragios. Performance vs computational efficiency for optimizing single and dynamic mrfs: Setting the state of the art with primal-dual strategies. *Computer Vision and Image Understanding*, 112(1):14–29, 2008.

[9] S. Larjavaara, R. Mantyla, T. Salminen, H. Haapasalo, J. Raitanen, J. Jaaskelainen, and A. Auvinen. Incidence of gliomas by anatomic location. *Neuro-oncology*, 9:319–325, Jul 2007.

[10] F. Michel, M. Bronstein, A. Bronstein, and N. Paragios. Boosted metric learning for 3d multi-modal deformable registration. In *ISBI*, pages 1209–1214, 2011.

[11] N. Moon, E. Bullitt, K. V. Leemput, and G. Gerig. Model-based brain and tumor segmentation. In *ICPR (1)*, pages 528–531, 2002.

[12] L. G. Nyul and J. K. Udupa. On standardizing the mr image intensity scale. *Magnetic Resonnance in Medicine*, 42:1072–1081, Dec 1999.

[13] S. Parisot, H. Duffau, S. Chemouny, and N. Paragios. Graph based spatial position mapping of low-grade gliomas. In *MICCAI (2)*, pages 508–515, 2011.

[14] L. K. O. Pokrajac, Megalooikonomou. Applying spatial distribution analysis techniques to classification of 3d medical images. *Artificial Intelligence in Medicine*, 33(3):261–280, 2005.

[15] M. Prastawa, E. Bullitt, S. Ho, and G. Gerig. A brain tumor segmentation framework based on outlier detection. *Medical Image Analysis*, 8(3):275–283, 2004.

[16] P. J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

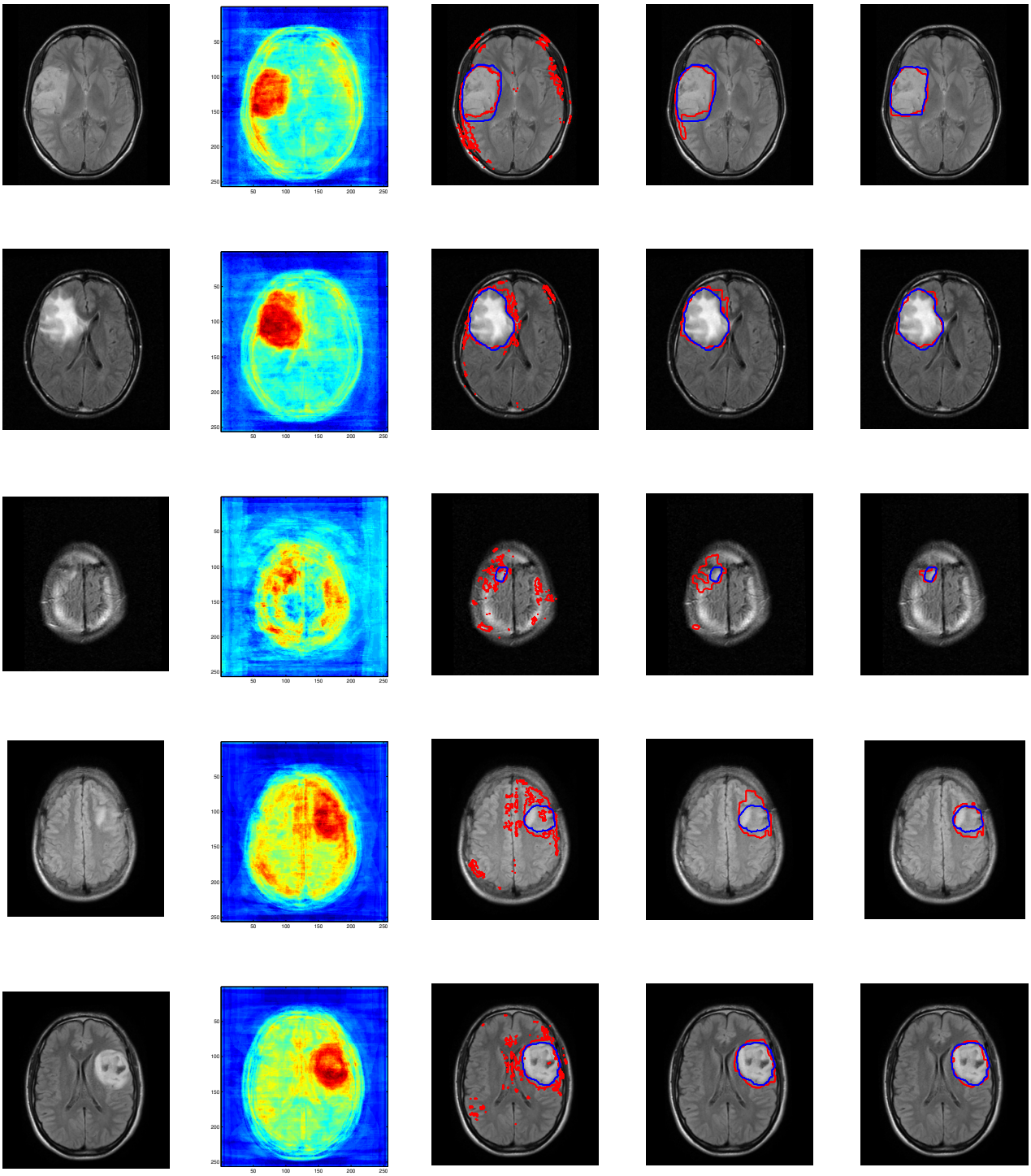


Figure 5: Various stages of segmentation. Original MRI image (a), Boosting score (b), Segmentation using boosting only (c), pairwise regularization (d) and spatial prior (e). The manual segmentation map (blue) is superimposed to the automatic segmentation (red)