Segmentation of Brain Tumor Images Based on Integrated Hierarchical Classification and Regularization

Stefan Bauer[1], Thomas Fejes[1,2], Johannes Slotboom[2], Roland Wiest[2], Lutz-P. Nolte[1], and Mauricio Reyes[1]

[1] Institute for Surgical Technology and Biomechanics, University of Bern
[2] Inselspital, Bern University Hospital, Switzerland
stefan.bauer@istb.unibe.ch

Abstract. We propose a fully automatic method for brain tumor segmentation, which integrates random forest classification with hierarchical conditional random field regularization in an energy minimization scheme. It has been evaluated on the BRATS2012 dataset, which contains low- and high-grade gliomas from simulated and real-patient images. The method achieved convincing results (average Dice coefficient: 0.73 and 0.59 for tumor and edema respectively) within a reasonably fast computation time (approximately 4 to 12 minutes).

1 Introduction

Fast and accurate segmentation of brain tumor images is an important but difficult task in many clinical applications. In recent years, a number of different automatic approaches have been proposed [1], but despite significant intra- and inter-rater variabilities and the large time consumption of manual segmentation, none of the automatic approaches is in routine clinical use yet. However, with the anticipated shift from diameter-based criteria to volume-based criteria in neuroradiological brain tumor assessment, this is likely to change in the future.

We are presenting a fully automatic method for brain tumor segmentation, which is based on classification with integrated hierarchical regularization. Not only does it offer to separate healthy from pathologic tissues, but it also subcategorizes the healthy tissues into CSF, WM, GM and the pathologic tissues into necrotic, active and edema compartment.

2 Methods

The general idea is based on a previous approach presented in [2]. After preprocessing (denoising, bias-field correction, rescaling and histogram matching) [6], the segmentation task is modeled as an energy minimization problem in a conditional random field (CRF) [8] formulation. The energy consists of the sum of the singleton potentials in the first term and the pairwise potentials in the second

term of equation (1). The expression is minimized using [7] in a hierarchical way similar to [2].

$$E = \sum_i V(y_i, \mathbf{x}_i) + \sum_{ij} W(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) \quad (1)$$

The singleton potentials $V(y_i, \mathbf{x}_i)$ are computed according to equation (2), where y_i is the label output from the classifier, $\mathbf{x}_i$ is the feature vector and δ is the Kronecker-δ function.

$$V(y_i, \mathbf{x}_i) = p(y_i|\mathbf{x_i}) \cdot (1 - \delta(\tilde{y}_i, y_i)) \quad (2)$$

In contrast to our previous approach, here we make use of random forests [4], [3] as a classifier instead of support vector machines (SVM). Random forests are ensembles of decision trees, which are randomly different. Training on each decision tree is performed by optimizing the parameters of a split function at every tree node via maximizing the information gain when splitting the training data. For testing, the feature vector is pushed through each tree, applying a test at each split node until a leaf node is reached. The label posterior is calculated by averaging the posteriors of the leave nodes from all trees $p(y_i|\mathbf{x}_i) = 1/T \cdot \sum_t^T p_t(y_i|\mathbf{x}_i)$. Compared to SVMs, random forests have the advantage of being able to naturally handle multi-class problems and they provide a probabilistic output instead of hard label separations [5]. We use the probabilistic output for the weighting factor $p(y_i|\mathbf{x}_i)$ in equation (2), in order to control the degree of spatial regularization based on the posterior probability of each voxel label. A 28-dimensional feature vector is used for the classifier, which combines the intensities in each modality with the first-order textures (mean, variance, skewness, kurtosis, energy, entropy) computed from local patches around every voxel in each modality.

We have also developed an improved way to compute the pairwise potentials $W(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j)$, which account for the spatial regularization. In equation (3) $w_s(i, j)$ is a weighting function, which depends on the voxel spacing in each dimension. The term $(1 - \delta(y_i, y_j))$ penalizes different labels of adjacent voxels, while the intensity term $\exp\left(\frac{\mathrm{PCD}(\mathbf{x}_i - \mathbf{x}_j)}{2 \cdot \bar{x}}\right)$ regulates the degree of smoothing based on the local intensity variation, where PCD is a pseudo-Chebyshev distance and $\bar{x}$ is a generalized mean intensity. $D_{pq}(y_i, y_j)$ allows us to incorporate prior knowledge by penalizing different tissue adjancencies individually.

$$W(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) = w_s(i, j) \cdot (1 - \delta(y_i, y_j)) \cdot \exp\left(\frac{\mathrm{PCD}(\mathbf{x}_i - \mathbf{x}_j)}{2 \cdot \bar{x}}\right) \cdot D_{pq}(y_i, y_j) \quad (3)$$

3 Results

The performance of the proposed method has been evaluated on the BRATS2012 dataset [3] using 5-fold cross-validation. The BRATS2012 dataset contains skull-

[3] http://www2.imm.dtu.dk/projects/BRATS2012/

stripped multimodal MR images (T_1, $T_{1contrast}$, T_2, Flair) of 80 low- and high-grade gliomas from simulations and real patient cases (1mm isotropic resolution). In order to be compatible with the BRATS ground truth, our "necrotic" and "active" labels were combined to form the "core" label, the "edema" label was unmodified and all other labels were ignored.

Quantitative results for different overlap and surface distance metrics, which were obtained using the BRATS2012 online evaluation tool, are detailed in table 1 and exemplary image results are shown in figure 1. Computation time for the segmentation ranged from 4 to 12 minutes depending on the size of the dataset.

We also compared the proposed approach to our previous method [2] which used SVMs as a classifier instead of random forests and which had a less sophisticated regularization. With the new method, the computation time could be reduced by more than a factor of two and the accuracy measured by the Dice coefficient was also improved.

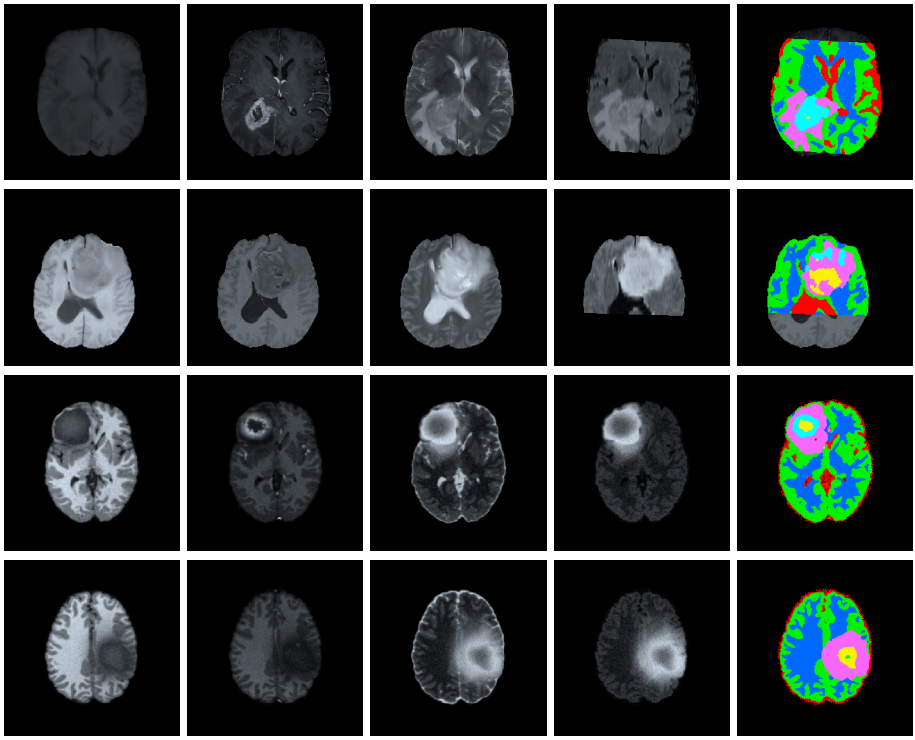

Fig. 1. Exemplary image results shown on one axial slice for a high-grade glioma patient (first row), a low-grade glioma patient (second row), a simulated high-grade glioma dataset (third row) and a simulated low-grade glioma dataset (last row). Each row shows from left to right: T_1, $T_{1contrast}$, T_2, Flair image and the label map obtained from the automatic segmentation (color code: red=CSF, green=GM, blue=WM, yellow=necrotic, turquoise=active, pink=edema.)

Table 1. Quantitative results from the BRATS2012 online evaluation tool. HG stands for high-grade, LG for low-grade and Sim for the simulated glioma datasets. The metrics in the table from left to right are: Dice, Jaccard, sensitivity, specificity, average distance, Hausdorff distance, Cohen's kappa.

		Dice	Jaccard	Sens.	Spec.	AD [mm]	HD [mm]	Kappa
HG	edema	0.61±0.15	0.45±0.15	1.0±0.0	0.56±0.15	5.0±5.3	60±31	0.32±0.25
	tumor	0.62±0.27	0.50±0.25	1.0±0.0	0.59±0.31	6.3±7.8	69±25	
LG	edema	0.35±0.18	0.23±0.13	1.0±0.0	0.49±0.23	10.4±9.2	69±28	0.07±0.23
	tumor	0.49±0.26	0.36±0.24	1.0±0.0	0.49±0.28	5.4±3.8	53±32	
Sim-HG	edema	0.68±0.26	0.56±0.26	1.0±0.0	0.90±0.07	1.3±0.7	12±6	0.67±0.13
	tumor	0.90±0.06	0.81±0.09	1.0±0.0	0.91±0.08	1.5±1.7	16±10	
Sim-LG	edema	0.57±0.24	0.44±0.22	1.0±0.0	0.84±0.17	1.6±0.9	10±6	0.38±0.18
	tumor	0.74±0.10	0.59±0.12	1.0±0.0	0.77±0.18	2.6±1.1	16±5	
All	edema	0.59±0.24	0.45±0.23	1.0±0.0	0.75±0.22	3.5±5.1	30±31	0.42±0.27
	tumor	0.73±0.22	0.61±0.23	1.0±0.0	0.73±0.26	3.6±4.6	34±29	

4 Discussion and Conclusion

We have presented a method for fully automatic segmentation of brain tumors, which achieves convincing results within a reasonable computation time on clinical and simulated multimodal MR images. Thanks to the ability of the approach to delineate subcompartments of healthy and pathologic tissues, it can have a significant impact in clinical applications, especially tumor volumetry. To evaluate this more thoroughly, a prototype of the method is currently being integrated into the neuroradiology workflow at Inselspital, Bern University Hospital.

References

1. Angelini, E.D., Clatz, O., Mandonnet, E., Konukoglu, E., Capelle, L., Duffau, H.: Glioma Dynamics and Computational Models: A Review of Segmentation, Registration, and In Silico Growth Algorithms and their Clinical Applications. Current Medical Imaging Reviews 3(4) (2007)
2. Bauer, S., Nolte, L.P., Reyes, M.: Fully automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization. In: MICCAI. LNCS, vol. 14. Springer, Toronto (2011)
3. Bochkanov, S., Bystritsky, V.: ALGLIB, `www.alglib.net`
4. Breiman, L.: Random forests. Machine Learning 45(1) (2001)
5. Criminisi, A., Shotton, J., Konukoglu, E.: Decision Forests for Classification , Regression , Density Estimation , Manifold Learning and Semi-Supervised Learning. Tech. rep., Microsoft Research (2011)
6. Ibanez, L., Schroeder, W., Ng, L., Cates, J., Others: The ITK software guide (2003)
7. Komodakis, N., Tziritas, G., Paragios, N.: 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) (2008)
8. Lafferty, J., McCallum, A., Pereira, F.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: ICML Proceedings. Citeseer (2001)