Fully Automatic Segmentation of Brain Tumor Images Using Support Vector Machine Classification in Combination with Hierarchical Conditional Random Field Regularization

Stefan Bauer[*], Lutz-P. Nolte, and Mauricio Reyes

Institute for Surgical Technology and Biomechanics,
University of Bern, Switzerland
`stefan.bauer@istb.unibe.ch`

Abstract. Delineating brain tumor boundaries from magnetic resonance images is an essential task for the analysis of brain cancer. We propose a fully automatic method for brain tissue segmentation, which combines Support Vector Machine classification using multispectral intensities and textures with subsequent hierarchical regularization based on Conditional Random Fields. The CRF regularization introduces spatial constraints to the powerful SVM classification, which assumes voxels to be independent from their neighbors. The approach first separates healthy and tumor tissue before both regions are subclassified into cerebrospinal fluid, white matter, gray matter and necrotic, active, edema region respectively in a novel hierarchical way. The hierarchical approach adds robustness and speed by allowing to apply different levels of regularization at different stages. The method is fast and tailored to standard clinical acquisition protocols. It was assessed on 10 multispectral patient datasets with results outperforming previous methods in terms of segmentation detail and computation times.

Keywords: MRI, Segmentation, Brain Tumor, Glioma, SVM, CRF.

1 Introduction

Segmentation of healthy and pathologic brain tissues from magnetic resonance images (MRI), including their subregions, is important in cancer treatment planning as well as for cancer research. In current clinical practice, the analysis of brain tumor images is mostly done manually. Apart from being time-consuming, this has the additional drawback of significant intra- and interrater variability, which was reported to be around 20%±15% and 28%±12% respectively according to Mazzara et al. [8].

Nowadays, in clinical practice, usually four different MRI modalities are used to delineate the tumor and its subregions from MRI images of the head. These

---

modalities include T_1-weighted images (T_1), T_1-weighted images with contrast enhancement (T_{1c}), T_2-weighted images (T_2) and $T_{2-\text{flair}}$-weighted images (T_{2f}). The tumor area is usually divided into necrotic, active and edema subregion. Each of the modalities reveal different subregions. In general, the radiologist considers all these MRI modalities simultaneously when segmenting brain tumor images. Standard clinical acquisition protocols for brain tumor imaging provide high intra-slice resolution, but poor inter-slice spacing. This poses an additional challenge for automatic analysis. In order to cope with all these challenges, we developed a new method, which is clinically-oriented and does not only make use of sophisticated image-processing methods, but also takes the type of data into account, which is usually available for large-scale clinical studies and patient treatment planning.

In this study, we focus on the case of gliomas, which is the most aggressive type of brain tumors. We treat the problem at hand as a classification task that maps each voxel to its corresponding label based on a multidimensional feature vector. Machine learning techniques have proven to yield good results in many cases. For example Verma et al. [11] use SVMs to classify brain tumors from a large number of different MRI modalities. However, these methods assume that the data is independent and identically distributed (iid), which is clearly not the case for image voxels. Most voxel labels strongly depend on their neighbors. These spatial relationships can be effectively considered using Conditional Random Field (CRF) methods. We suggest to use the CRF method to regularize the result produced by the classifier retrospectively, in a way inspired by the approach of Lee et al. [6]. The classifier is first trained under the assumption that the data is iid and the result is regularized subsequently using a CRF approach. This makes the whole process very efficient while maintaining improved accuracy by considering neighbor relationships.

In order to exploit the full potential of the classification method, we propose not only to segment the brain into healthy and tumor regions as it is usually done, but also subdivide the healthy region into cerebrospinal fluid (CSF), gray matter (GM) and white matter (WM) and subdivide the tumor area into necrotic part, active part and edema part. We suggest to employ a novel hierarchical approach for this task in order to increase robustness and speed. The hierarchical approach allows us to apply a strong regularization after segmenting the brain into tumor and healthy regions. Each of the two initially identified regions is then classified into its respective subregions before a weaker regularization is applied to the labeled image of the healthy and tumor subregions.

2 Methods

2.1 Feature Extraction

Features are extracted from the multispectral imaging data. The most prominent features for distinguishing pathologic and healthy tissues, as well as all their subregions, are the image intensities in the different modalities. Additionally, we use the first order texture features (mean, variance, skewness, kurtosis, energy,

entropy) according to [10]. First-order textures can be computed fast and easily from small patches around each voxel in all four modalities. This yields a 28-dimensional feature vector $\mathbf{x}$, which consists of the voxel-wise concatenation of the multimodal intensities I and multimodal textures $\mathbf{T}$ at each voxel i as shown in equation (1).

$$\mathbf{x}(i) = [I_{T_1}(i), I_{T_{1c}}(i), I_{T_2}(i), I_{T_{2f}}(i), \mathbf{T}_{T_1}(i), \mathbf{T}_{T_{1c}}(i), \mathbf{T}_{T_2}(i), \mathbf{T}_{T_{2f}}(i)] \quad (1)$$

2.2 Classification

Classification is done using a soft-margin SVM classifier [9]. SVMs are discriminative classifiers, originating from machine learning. They require a training step to find a separating hyperplane for the data in the feature space. SVMs solve the optimization problem

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^{l} \xi_i \quad (2)$$

$$\text{subject to } y_i(\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \text{ with } \xi_i \geq 0 \quad (3)$$

where $(\mathbf{x}_i, y_i)$ are the instance-label pairs of the dataset and $\mathbf{w}$ is the normal vector of the separating hyperplane. C is a penalty parameter for the error term and b is the offset of the hyperplane. The appealing property of SVMs is that they offer the possibility to use a kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ for transforming the data into a higher-dimensional feature space, where the data can be linearly separated efficiently with a maximum margin. Slack variables ξ_i are used for soft-margin classification. Parameter selection for the SVM classifier with a radial basis function (RBF) kernel is done using grid-based cross-validation on the training data. In order to extend the inherently binary SVM classifier to a multiclass problem, we use a one-against-one voting strategy. Our SVM classification is based on the LibSVM implementation described in [2].

2.3 Regularization

Regularization is done in two different stages using a CRF method. Conditional Random Fields are an extension of Markov Random Fields (MRF). Both offer the possibility to formulate the regularization as an energy minimization problem. To this end we use a second-order CRF with two energy terms.

$$E = \sum_i V(y_i) + \sum_{ij} w(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) D_{pq}(y_i, y_j) \quad (4)$$

The first term in equation (4) denotes the data energy, which is only dependent on the data at the current point, whereas the second term constitutes the smoothness energy, which takes the neighborhood information into account. For the case at hand, V is the unary potential function, w describes the neighborhood relationships and D_{pq} is a label distance function for the pairwise potentials. The unary potentials can be calculated directly from the voxel-wise output produced

by the SVM classifier. The smoothness energy is computed depending on the neighboring voxels.

We compute the unary potentials $V(y_i)$ from the label output y_i of the SVM classifier. In equation (5), c_1 is a constant which allows us to adjust the weights of unary and pairwise potentials. $\tilde{y}_i$ is the newly assigned label after regularization and δ is the Kronecker δ function.

$$V(y_i) = c_1 \cdot (1 - \delta(\tilde{y}_i, y_i)) \tag{5}$$

For computing the pairwise potentials, we suggest a new formulation, which is an extension of the approach given by Boykov et al. [1]. In equation (6), $c_2(i,j)$ is a weighting function, which allows us to apply different weights for different neighbors. The term $(1 - \delta(y_i, y_j))$ is the most important, penalizing different labels y of adjacent voxels. The last term in equation (6) imposes strong smoothness constraints in regions of similar intensities $\mathbf{x}$ and relaxes regularization in regions of high intensity contrast. The intensity term is adapted for multimodal images and $\bar{x}$ is a generalized multimodal mean intensity. The label distance function D_{pq} in equation (4) is used to penalize adjacencies of necrotic or active tumor regions with healthy tissues more strongly because these adjacencies are less likely to occur.

$$w(y_i, y_j, \mathbf{x}_i, \mathbf{x}_j) = c_2(i,j) \cdot (1 - \delta(y_i, y_j)) \cdot \exp\left(\frac{||\mathbf{x}_i - \mathbf{x}_j||}{2 \cdot \bar{x}}\right) \tag{6}$$

As mentioned before, the regularization is carried out in a hierarchical way. After the first coarse classification into tumor and healthy tissues, a strong 3D regularization is employed using a von Neumann neighborhood (6-neighboring voxels in 3D). However, in our approach costs w_{ij} in z-direction are weighted with a smaller constant $c_2(i,j)$ because the z-spacing of slices is much larger than the in-slice-spacing. If the same weight would be applied to all neighbors, the label map would be oversmoothed in z-direction.

In a second stage, regularization is applied on the image, which has been subclassified into the different tumor and healthy subregions. In this case very fine structures, especially for CSF, are present. These fine structures do not allow to use a 3D regularization. Due to the large z-spacing, the fine structures would be oversmoothed. Therefore, we use only a 2D neighborhood for the second stage, but with a Moore neighborhood structure now (8 neighbors in 2D), which renders smoother results.

Optimizing CRFs is a challenging and computationally intensive task, especially when many voxels and multiple labels are involved. We employ a recent optimization algorithm proposed by Komodakis et al. [5], which is based on linear programming via graph-cuts and primal-dual decomposition. The algorithm is able to efficiently derive almost optimal solutions for a wide class of NP-hard CRF problems. The most important reason for us to use this optimization approach is its great computational efficiency, which renders it much faster than conventional graph-cuts optimization techniques.

2.4   Application to Brain Tumor Image Analysis

Initially, the images undergo a preprocessing pipeline. In a first step, the four modalities are registered with the help of a a rigid registration and mutual information metric. Next, only the brain region is extracted from the images using a fully-automatic, customized skull-stripping algorithm[1]. Subsequently, noise is removed with an edge-preserving smoothing filter and the bias field is corrected. In order to allow for interpatient classification, the image histograms are matched across all patients for each MRI modality separately. The preprocessing is completely integrated with the SVM classification and CRF regularization components using the Insight Toolkit for Segmentation and Registration (ITK) [4].

3   Results

3.1   Image Data

We evaluated our algorithm on images of 10 patients from the ContraCancrum brain tumor database [7]. Each patient dataset consists of a T_1, T_{1c}, T_2 and T_{2f} image. The images were resampled to an in-slice resolution of 1mm (210x210 to 260x260 voxels), the inter-slice spacing was between 3mm and 6mm depending on the dataset (19 to 25 slices). Figure 1 shows an axial slice of one patient with the four MRI modalities we used.
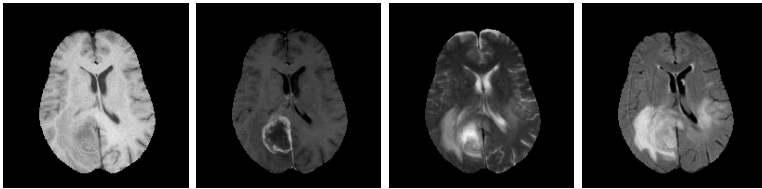

Fig. 1. Axial slice of one patient. The four MRI modalities under study are shown, from left to right: the T_1, T_{1c}, T_2 and T_{2f} image.

3.2   Evaluation

For quantitative evaluation of the results we use the Dice similarity coefficient. The ground truth was defined by a manual segmentation. However, the manual segmentation was only available for the tumor tissues, not for the healthy tissues. Therefore only the accuracy of the tumor segmentation could be evaluated quantitatively. The accuracy of the segmentation of healthy tissues could only be rated qualitatively by visual inspection.

We subdivided our results into intra- and interpatient case. For the intrapatient case, the classifier was trained on small subregions of the relevant areas

---

[1] available from:
http://www.istb.unibe.ch/content/surgical_technologies/
medical_image_analysis/software/

in the same patient, while in the interpatient case we performed leave-one-out cross-validation. Additionally, in both cases, we compared our proposed hierarchical, regularized approach with a non-hierarchical classification, that does not comprise the proposed two-step CRF-regularization.

Figure 2 presents the segmentation result in case of interpatient training using leave-one-out cross-validation for the same slice as shown in figure 1. The gross tumor volume (GTV), comprising all tumor tissues, is well delineated, although the segmentation appears noisy if no regularization is applied. Classifying the tumor subregions appears to be more challenging, but our proposed approach performs still better than the method without hierarchy and regularization. From a visual inspection, the segmentation of healthy tissues (CSF,
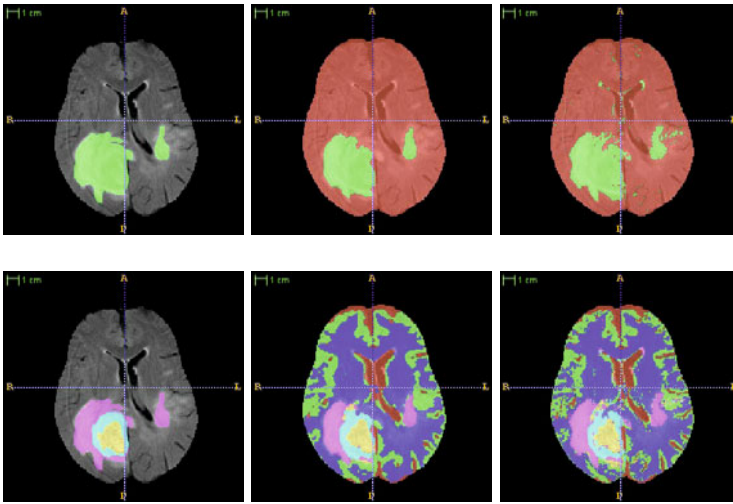

Fig. 2. Results for one axial slice of the interpatient leave-one-out case. First row: coarse classification into tumor and healthy tissues, second row: fine classification into necrotic, active, edema part and CSF, GM, WM respectively. From left to right: manual segmentation, hierarchical SVM-classification with CRF-regularization, non-hierarchical SVM-classification without regularization.

Table 1. Dice similarity coefficient (mean and standard deviation) for the ten datasets under study. Gross tumor volume (GTV) comprises the complete tumor region, including active, necrotic and edema part. Results are shown for the intra- and interpatient (leave-one-out cross-validation) case, with and without regularization.

	GTV	Necrotic	Active	Edema
Intrapatient Regularized	$\mathbf{0.84 \pm 0.03}$	0.61 ± 0.24	0.71 ± 0.09	0.73 ± 0.04
Intrapatient Unregularized	$\mathbf{0.76 \pm 0.10}$	0.45 ± 0.31	0.59 ± 0.16	0.72 ± 0.07
Interpatient Regularized	$\mathbf{0.77 \pm 0.09}$	0.45 ± 0.23	0.64 ± 0.13	0.60 ± 0.16
Interpatient Unregularized	$\mathbf{0.67 \pm 0.13}$	0.30 ± 0.24	0.46 ± 0.12	0.63 ± 0.16

GM, WM) appears to be reasonable in all cases, but again the result has many outliers if no regularization is applied.

The quantitative results using Dice similarity coefficient are summarized in table 1. With hierarchical regularization, the mean Dice similarity coefficient over all 10 patients is 0.84 for the intrapatient case and 0.77 for the interpatient leave-one-out case when the GTV is considered. Dice coefficients for the individual tumor subregions are lower. However, the results show once again that the Dice coefficient is clearly worse when no hierarchical regularization is applied.

Computation time for the segmentation algorithm (excluding preprocessing) is between 20 and 120 seconds on a single CPU running at 2.33 GHz. Computation time mainly depends on the size of the image dataset and on the complexity of the SVM optimization.

4   Discussion and Conclusion

We presented a clinically-oriented method to segment 3D MRI images of brain tumor patients into healthy and tumor areas, including their individual subregions. For this, we propose to apply a hierarchical SVM-based classification and combine it with a CRF-based regularization in two stages.

Dice similarity coefficients for the GTV are in a similar range as the results presented in [3] and [12], but different data was used. In contrast to those approaches, our method additionally returns all the tumor and healthy subregions while being faster in computation time. The accuracy of our automatic method lies in a similar range as the values reported for inter-observer variability of manual segmentations in [8]. However, the automatic method has advantages in longitudinal studies because the results are not biased subjectively. Due to the additional difficulty in subdividing the tumor region, Dice similarity coefficients for the individual tumor subregions are lower than for the gross tumor volume.

We show that the combination of SVM and CRF methods offers two complementary strengths. The results mentioned in table 1 suggest that our approach including hierarchical regularization improves the overlap with the ground truth significantly. This regularization seems to be even more important when training is applied on different patients. With the hierarchical regularization, the Dice coefficients are only slightly worse in interpatient training compared to intrapatient training. This is an important finding because the possibility to perform training on a disjoint subset of patients is crucial for efficient segmentation procedures in clinical practice. A visual inspection of the images in figure 2 shows that the segmentation results for healthy tissues as well as for tumor tissues appear much noisier when no CRF regularization is applied and inconsistencies of adjacent slices can be observed.

In contrast to other methods, our proposed segmentation approach yields the tumor and healthy region, as well as all their subregions. It adds powerful spatial regularization to iid classification and runs fast on clinically relevant data.

References

1. Boykov, Y., Jolly, M.: Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. In: Proc. ICCV, pp. 105–112. IEEE, Los Alamitos (2002)
2. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines (2001), http://www.csie.ntu.edu.tw/~cjlin/libsvm
3. Corso, J.J., Sharon, E., Dube, S., El-Saden, S., Sinha, U., Yuille, A.: Efficient multilevel brain tumor segmentation with integrated bayesian model classification. IEEE T. Med. Imaging 27(5), 629–640 (2008)
4. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: The ITK Software Guide. Kitware, 2nd edn. (2005), http://www.itk.org/ItkSoftwareGuide.pdf
5. Komodakis, N., Tziritas, G., Paragios, N.: gle and dynamic MRFs: Setting the state of the art with primal-dual strategies. Comput. Vis. Image Und. 112(1), 14–29 (2008)
6. Lee, C., Wang, S., Murtha, A., Brown, M., Greiner, R.: Segmenting Brain Tumors using Pseudo–Conditional Random Fields. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 359–366. Springer, Heidelberg (2008)
7. Marias, K., Dionysiou, D., Sakkalis, V., Graf, N., Bohle, R.M., Coveney, P.V., Wan, S., Folarin, A., Buechler, P., Reyes, M., Clapworthy, G., Liu, E., Sabczynski, J., Bily, T., Roniotis, A., Tsiknakis, M., Giatili, S., Veith, C., Messe, E., Stenzhorn, H., Kim, Y.j., Zasada, S., Haidar, A.N., Bauer, S., Wang, T., Zhao, Y., Karasek, M., Grewer, R., Franz, A., Stamatakos, G.: Clinically driven design of multi-scale cancer models: the ContraCancrum project paradigm. J. Roy. Soc. Interface Focus 1(3), 450–461 (2011)
8. Mazzara, G., Velthuizen, R., Pearlman, J., Greenberg, H., Wagner, H.: Brain tumor target volume determination for radiation treatment planning through automated MRI segmentation. Int. J. Radiat. Oncol., Biol. Phys. 59(1), 300–312 (2004)
9. Schoelkopf, B., Smola, A.: Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge (2002)
10. Tuceryan, M., Jain, A.: Texture analysis, 2nd edn., pp. 207–248. World Scientific Publishing, Singapore (1998)
11. Verma, R., Zacharaki, E., Ou, Y., Cai, H., Chawla, S., Lee, S., Melhem, E., Wolf, R., Davatzikos, C.: Multiparametric Tissue Characterization of Brain Neoplasms and Their Recurrence Using Pattern Classification of MR Images. Acad. Radiol. 15(8), 966–977 (2008)
12. Wels, M., Carneiro, G., Aplas, A., Huber, M., Hornegger, J., Comaniciu, D.: A Discriminative Model-Constrained Graph Cuts Approach to Fully Automated Pediatric Brain Tumor Segmentation in 3-D MRI. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part I. LNCS, vol. 5241, pp. 67–75. Springer, Heidelberg (2008)