A Boosting Approach for Multiple Sclerosis Lesion Segmentation in Multi-Spectral 3D MRI

Michael Wels[1], Martin Huber[2], and Joachim Hornegger[1]

[1] Institute of Pattern Recognition, University Erlangen-Nuremberg, Germany,
[2] Siemens AG, CT SE SCR 2, Erlangen, Germany.

Abstract. In this paper we present a fully automated approach to multiple sclerosis lesion segmentation in multi-spectral brain MRI data. The proposed segmentation framework is based on the recently introduced probabilistic boosting trees, which is a strategy for supervised learning. Foreground and background voxels are distinguished by considering a 2D context surrounding the voxel of interest and its transformation to a large set of Haar-like features. This allows for building a discriminative model that captures class specific characteristics despite of the well-known drawbacks of MR imaging. Training the model itself involves two major steps: successively selecting and combining features that best separate the training data by means of boosting and inductively grouping the resulting boosted classifiers in a tree structure. By applying the equivalence of boosting and additive logistic regression the approach is capable to derive a discriminative model for voxel classification in terms of posterior probabilities. The final segmentation is obtained after post-processing the preliminary result by stochastic relaxation and a standard zero level set segmentation approach. The applicability of the proposed method is demonstrated by quantitative evaluation within a leave-one-patient-out cross-validation.

1 Introduction

Quantitatively assessing disease progression is a major concern in the case of multiple sclerosis (MS)—a common neuropathological disease of young adults that primarily affects cerebral white matter within the human brain. One of the indices characteristic for the progression of the disease is lesion volume. However, manual quantitative measurements on volumetric magnetic resonance (MR) images suffer from substantial intra- and inter-rater variability [1], such that providing accurate and reliable automatic segmentation tools to facilitate valid MS lesion quantification is of enormous interest. Therefore, we propose a knowledge-driven approach for MS lesion segmentation in multi-spectral 3D brain MRI data based on the recently introduced [2] probabilistic boosting trees (PBT).

2 Methods

In order to distinguish foreground, i.e. lesion, and background, i.e. non-lesion, voxels within multi-spectral (FLAIR, T1, T2) MR volumetric data ($408 \times 512 \times 19$

and $408 \times 512 \times 21$) without contrast enhancement we build a PBT in a similar manner to [3] to derive a discriminative model in terms of posterior probabilities for individual voxels from manually segmented training data. Subsequently in the segmentation process, the results obtained by PBT are refined by stochastic relaxation [4] and a standard zero level set approach from the Insight Segmentation and Registration Toolkit (ITK, www.itk.org). The latter uses anisotropic diffusion filtering [5] on one of the input images to obtain a feature image guiding evolution of the zero level set. As the axial resolution of the input data is low we restrict our approach to operate on down-sampled 2D axial slices (256×256) without taking into account inter-slice voxel neighborhood relations. The overall processing pipeline is depicted in Fig. 1.

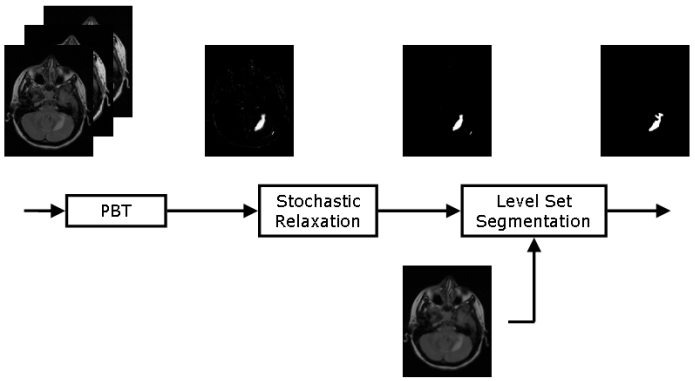


Fig. 1. The proposed segmentation framework.

The framework's first step PBT recursively groups boosted ensembles of weak classifiers to a tree structure during learning from annotated data. When Discrete AdaBoost [6] is chosen as the boosting strategy this resembles building a binary regression tree as the final boosted classifier

$$H(\boldsymbol{x}) = \sum_{t=1}^{T} \alpha_t h_t(\boldsymbol{x}) \tag{1}$$

generated within each inner node for a feature vector $\boldsymbol{x}$ through a weighted combination of $T \in \mathbb{N}$ weak classifiers $h_t(\boldsymbol{x})$ with individual weights α_t, $t \in \{\, 1, \ldots, T \,\}$, asymptotically approaches the additive logistic regression model [7]:

$$H(\boldsymbol{x}) \approx \frac{1}{2} \ln \frac{p(y = 1|\boldsymbol{x})}{p(y = -1|\boldsymbol{x})} \tag{2}$$

where $y \in \{\, -1, 1 \,\}$ denotes the classification outcome. Therefore, at each inner node v of the resulting PBT with strong classifier H_v and outgoing arrows r_v^{-1}

and r_v^1 associated with the possible classifications an approximation of the overall posterior probability $\tilde{p}_v(y|\boldsymbol{x})$ can be computed via the recursive formula

$$\begin{aligned}\tilde{p}_v(y|\boldsymbol{x}) = \tilde{p}_{\beta(r_v^{-1})}(y|\boldsymbol{x}) \cdot \frac{e^{-2H_v(\boldsymbol{x})}}{1+e^{-2H_v(\boldsymbol{x})}} \\ +\tilde{p}_{\beta(r_v^1)}(y|\boldsymbol{x}) \cdot \frac{e^{2H_v(\boldsymbol{x})}}{1+e^{2H_v(\boldsymbol{x})}}\end{aligned} \tag{3}$$

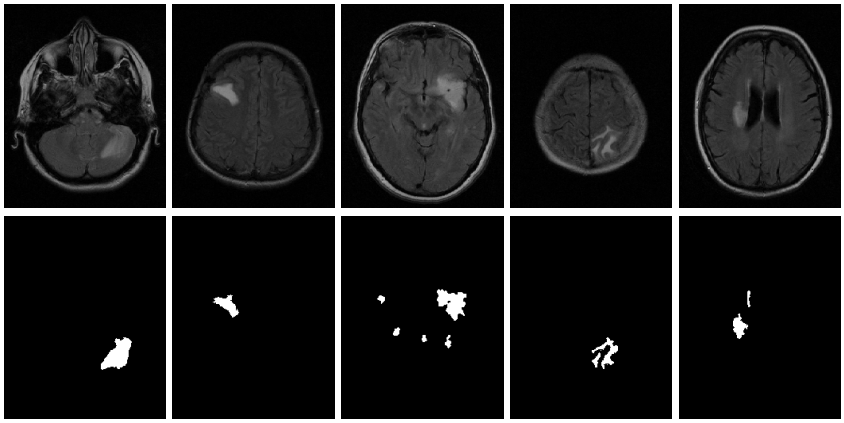
where $\beta(r)$ denotes the node where arrow r ends. At leaf nodes a hard classification $\tilde{p}_v(y=-1|\boldsymbol{x})=1$ and $\tilde{p}_v(y=1|\boldsymbol{x})=0$ or $\tilde{p}_v(y=-1|\boldsymbol{x})=0$ and $\tilde{p}_v(y=1|\boldsymbol{x})=1$ is returned.
For the purpose of classification it is tried to capture the structural variability of foreground and background voxels by not only considering multi-spectral intensity and gradient values but also, additionally, 17,472 Haar-like features [8] computed on a 15×15 square centered at the voxel of interest. Those features derived from a subset of the extended set of Haar-like feature prototypes [9] are represented implicitly in memory by so-called "Integral Images". This allows for fast re-computation of the features with respect to a given voxel when actually assessed.
For a given voxel i the posterior probabilities $p(y_i=1|\boldsymbol{x}_i)$ and $p(y_i=-1|\boldsymbol{x}_i)$ obtained by PBT that determine the segmentation are smoothed by stochastic relaxation independently from the initial features used by PBT itself. For this purpose the segmentation $\boldsymbol{y}$ is assumed to form a Markov random field (MRF) with individual spatial priors

$$p(y_i) = p(y_i|y_{\mathcal{N}_i}) = e^{-\frac{\beta}{2}\sum_{j\in\mathcal{N}_i} V_{ij}(y_i,\bar{y}_j)} \tag{4}$$

where $V_{ij}(y_i,\bar{y}_j)$ denotes the two-elemented clique potential of the classification y_i at voxel i and the mean classification $\bar{y}_j$ of a neighboring voxel j. In this notation $d(i,j)$ denotes the Euclidian distance. In our implementation we use 10 iterations of an algorithm similar to iterated conditional modes (ICM) [10] doing mean field-like approximation to the true posteriors $p(y_i)$. The neighborhood $\mathcal{N}_i$ considered for each individual voxel i is an intra-slice 8-neighborhood.

3 Material and Experimental Setting

For evaluation of the proposed method there were 6 manually segmented multi-spectral MRI scans (FLAIR, T1, T2) of sizes $408 \times 512 \times 21$ and $408 \times 512 \times 19$ available. It takes less than five minutes to process one of the MRI volumes in a non-optimized C++ implementation of our segmentation framework on a Fujitsu Siemens Computers notebook equipped with an Intel Pentium M 2.0 GHz processor and 2 GB of memory. In fact, processing may be significantly accelerated as soon as traversation of the learned PBT is properly restricted by soft thresholding as originally proposed in [3]. Though, we currently rely on complete traversation. The leave-one-patient-out approach was used to train



(a) (b) (c) (d) (e)

Fig. 2. Segmentation results obtained by leave-one-out validation. The first row shows selected slices of the FLAIR sequences of five different data sets. The second row shows the associated segmentation result

Table 1. Performance indices obtained by leave-one-out validation for all of the examined data sets.

	Dice	Pearson	Sens.	Spec.	PPV	NPV
1	0.7338	0.7356	0.8014	0.9989	0.6767	0.9994
2	0.7509	0.7575	0.6578	0.9995	0.8746	0.9984
3	0.5602	0.5601	0.5220	0.9987	0.6044	0.9981
4	0.8570	0.8581	0.9371	0.9967	0.7895	0.9992
5	0.0000	-0.0001	0.0000	0.9995	0.0000	0.9998
6	0.4912	0.5067	0.3929	0.9997	0.6550	0.9990

six different classifiers from approximately 70,000 randomly selected training samples, i.e. voxels inside the head of the patients, uniformly distributed over all the input slices. The maximum number of features selected by AdaBoost in each tree node were increased level-wise beginning with 1 at the root node. The maximum depth of the trees learned was restricted to 10. For stochastic relaxation $\beta = 1.2$ was chosen for empirical reasons. The settings for anisotropic diffusion filtering and Laplacian zero level set segmentation were adopted from ITK's introductory example. With the hardware and implementation mentioned above the duration of building one classification tree is about 12 hours.

4 Results

As can be seen from the quantitative results in Tab. 1 three of the six classifiers reach a Dice coefficient of more than 70% on their test data set, two reach about 50% and one fails with 0% due to the lack of significant MS lesions in the associated data volume. However, for a fully automated approach volumetric

overlap of more than 50% can be considered a remarkable achievement. Figure 2 gives a visual impression of the segmentation results obtained.

5 Conclusion

The method for segmentation of MS lesions in multi-spectral 3D brain MRI data discussed in this paper makes use of structural information by additionally taking into account the context of a voxel for the purpose of classification. The presented results show that by doing so supervised techniques like PBT can be employed for MRI tissue classification even though they are usually—when relying on individual voxel intensities only—considered inappropriate due to the typically large inter-scan variations. This encourages further investigation of medical image segmentation approaches based on boosting weak classifiers in the sense of features from large sets of feature candidates. Future work involves further assessment and refinement of the proposed method and investigation of alternative structural features that can be used in the context of boosting.

References

1. van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., Suetens, P.: Automated segmentation of multiple sclerosis lesions by model outlier detection. IEEE Trans. Med. Imag. **20**(8) (2001) 677–688
2. Tu, Z.: Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. In: IEEE Intl. Conf. Comput. Vis., Beijing, China. (2005) 1589–1596
3. Tu, Z., Zhou, X., Comaniciu, D., Bogoni, L.: A learning based approach for 3D segmentation and colon detagging. In: Europ. Conf. Comput. Vis., Graz, Austria. (2006) 436–448
4. Sonka, M., Fitzpatrick, J.M.: Handbook of Medical Imaging, Vol. 2: Medical Image Processing. SPIE—The International Society of Optical Engineering, Bellingham, Washington, USA (2004)
5. Weickert, J.: Anisotropic Diffusion in Image Processing. ECMI Series. Teubner-Verlag Stuttgart, Germany (1998)
6. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. In: Europ. Conf. Comput. Learning Theory. (1995) 23–37
7. Friedman, J., Hastie, T., Tibshirani, R.: Additive logistic regression: a statistical view of boosting. Stanford University, Department of Statistics, Technical Report (1998)
8. Viola, P., Jones, M.: Robust real-time object detection. In: 2nd International Workshop on Statistical Learning and Computational Theories of Vision, Vancouver, Canada. (2001) 1–25
9. Lienhart, R., Kuranov, A., Pisarevsky, V.: Empirical analysis of detection cascades of boosted classifiers for rapid object detection. In Michaelis, B., Krell, G., eds.: DAGM-Symposium. Volume 2781 of Lect. Notes in Comp. Sci., Springer (2003) 297–304
10. Besag, J.: On the statistical analysis of dirty pictures. J. R. Stat. Soc. Ser. B Stat. Methodol. **48**(3) (1986) 259–302