This is a preprint of the materials accepted for publication in ISSN 1054-6618, Pattern Recognition and Image Analysis, 2008, Vol. 18, No. 2, pp. 347-350. 

# Fully Automated Segmentation of Multiple Sclerosis Lesions in Multispectral MRI

M. Wels[1], M. Huber[2], J. Hornegger[1]

[1]University Erlangen-Nuremberg, Chair of Pattern Recognition, Martensstr. 3, 91058 Erlangen, Germany,
{michael.wels,joachim.hornegger}@informatik.uni-erlangen.de
[2]Siemens AG, CT SE SCR 2, Guenther-Scharowsky-Str. 1, 91058 Erlangen, Germany, martin.huber@siemens.com

This paper addresses segmentation of multiple sclerosis lesions in multispectral 3-D brain MRI data. For this purpose, we propose a novel fully automated segmentation framework based on probabilistic boosting trees, which is a recently introduced strategy for supervised learning. By using the context of a voxel to be classified and its transformation to an overcomplete set of Haar-like features, it is possible to capture class specific characteristics despite of the well-known drawbacks of MR imaging. By successively selecting and combining the most discriminative features during ensemble boosting within a tree structure, the overall procedure is able to learn a discriminative model for voxel classification in terms of posterior probabilities. The final segmentation is obtained after refining the preliminary result by stochastic relaxation and a standard level set approach. A quantitative evaluation within a leave-one-out validation shows
the applicability of the proposed method.

## Introduction

Quantitatively assessing disease progression is a major concern in the case of multiple sclerosis (MS)—a common disease of young adults that primarily affects cerebral white matter within the human brain. One of the indices characteristic of the evolution of the disease is the volume of the lesions that are observable on magnetic resonance (MR) images. However, quantitative measurements performed manually by radiologists suffer from substantial intra- and inter-rater variability [1], such that there is a need for accurate and reliable automatic segmentation tools to facilitate valid MS lesion quantification.

## Problem and Approach

For the purpose of MS lesion segmentation we aim to partition multispectral (FLAIR, T1, T2) MR volumetric data (408×512×19 and 408×512×21) without contrast enhancement into two regions—foreground, i.e. lesion, and background. In order to do so, we adopt a learning based approach called a probabilistic boosting tree (PBT) [2] in a similar manner to [3] to derive a discriminative model for individual image voxels from training data. Subsequently, the results obtained by PBT are refined by stochastic relaxation [4] and a standard zero level set approach that is implemented in 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. Due to the low axial

resolution of the input data, we restrict our approach to operate on down-sampled 2-D axial slices (256×256) without taking into account interslice voxel neighborhoods. Down-sampling is mainly due to the intent to capture ”macroscopic” context information by all the Haar-like features instead of local noise characteristics in the immediate neighborhood of each voxel. Figure 1 depicts the overall processing pipeline.

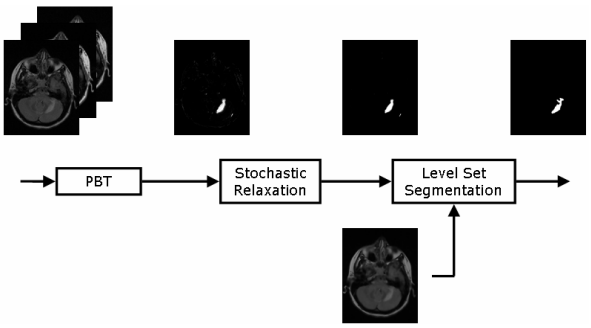



Figure 1: The proposed segmentation framework.

## Probabilistic boosting tree

As detailed in [2], PBT recursively groups boosted ensembles of weak hypothesis to a tree structure during learning from annotated training data. When used with Discrete AdaBoost [6], this resembles building a regression tree as the final hypothesis

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

generated for a feature vector $\boldsymbol{x}$ by Discrete AdaBoost through a weighted combination of $T \in \mathbb{N}$ weak hypothesis $h_t(\boldsymbol{x})$ with individual weights $\alpha_t$, $t \in \{1,\ldots,T\}$, asymptotically approaches an additive logistic regression model [7]:

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

where $y \in \{-1,1\}$ denotes the classification outcome. Therefore, at each inner vertex $v$ of the resulting PBT with strong hypothesis $H_v$ and outgoing arrows $r_{-1}$ and $r_1$ associated with the possible classifications, an approximation of the overall discriminative model $\tilde{p}_v(y \mid \boldsymbol{x})$ can be computed by the recursive formula

$$\tilde{p}_v(y \mid \boldsymbol{x}) = \tilde{p}_{\beta(r_{-1})}(y \mid \boldsymbol{x}) \cdot \frac{e^{-2H_v(\boldsymbol{x})}}{1+e^{-2H_v(\boldsymbol{x})}} + \tilde{p}_{\beta(r_1)}(y \mid \boldsymbol{x}) \cdot \frac{e^{2H_v(\boldsymbol{x})}}{1+e^{2H_v(\boldsymbol{x})}} \tag{3}$$

where $\beta(r)$ denotes the vertex where arrow $r$ ends. At leaf nodes, a hard classification $\tilde{p}_v(y=-1\mid\boldsymbol{x})=1$ and $\tilde{p}_v(y=1\mid\boldsymbol{x})=0$ or $\tilde{p}_v(y=-1\mid\boldsymbol{x})=0$ and $\tilde{p}_v(y=1\mid\boldsymbol{x})=1$ is returned.

## Haar-like Features

In order to adequately capture the structural variability of foreground and background voxels, the feature vectors $\boldsymbol{x}_i$ do not only consist of the multispectral intensity and gradient values at the associated index $i$ but also, additionally, of 17,472 Haar-like features 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 [8] are represented implicitly in memory by so-called “Integral Images”. This allows for fast recomputation of the features with respect to a given voxel when actually assessed. Details can also be found in [9].

## Stochastic Relaxation

The probabilities $p(y_i = 1 \mid \boldsymbol{x}_i)$ and $p(y_i = -1 \mid \boldsymbol{x}_i)$ obtained by the PBT are smoothed by stochastic relaxation assuming the segmentation **Y** to be a Markov random field with individual spatial priors

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

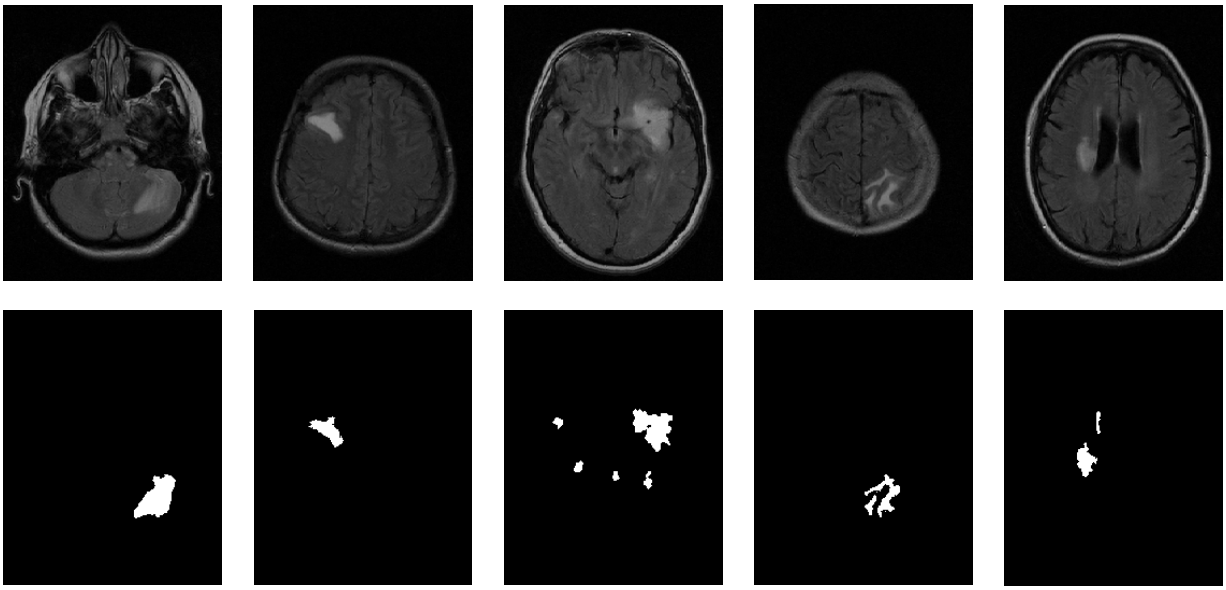
where $V_{ij}(y_i, \bar{y}_j)$ denotes the two-elemented clique potential of the classification $y_i$ at voxel $i$ and the mean classification $\bar{y}_j$ of a neighboring voxel $j$. We use 10 iterations of an algorithm similar to iterated conditional modes (ICM) [10] doing mean fieldlike approximation to the true posteriors $p(y_i)$. The neighborhood $N_i$ considered for each voxel $i$ is an intraslice 8-neighborhood.

As obvious edges tend to get oversmoothed by this approach, the final segmentation result for every slice is obtained by means of the Laplacian level set filter from ITK, which uses anisotropic diffusion filtering [5] on the slices of the FLAIR sequence to get a feature image guiding evolution of the zero level set (see Fig. 1).

## Material, Experimental Setting, and Results

For training and evaluation of the proposed method, there were 6 manually segmented multispectral MRI scans (FLAIR, T1, T2) of sizes 408×512×21 and 408×512×19 available. In a C++ implementation of our segmentation framework, it takes less than five minutes to process one of the MRI volumes

on a Fujitsu Siemens Computers notebook equipped with an Intel Pentium M 2.0 GHz processor and 2 GB of memory. The leave-one-out approach was used to train 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 were increased level wise beginning with 1. The maximum depth of the trees learned was restricted to 10. For stochastic relaxation, $\beta = 1.2$ was chosen for empirical reasons. The parameter settings for anisotropic diffusion filtering and Laplacian level set segmentation were adopted from ITK's introductory example. With the same hardware as above, building one classifier takes about 12 hours. The table depicts the quantitative results obtained when segmenting all of the available datasets. Results in terms of actual images are shown in Fig. 2. Although the visual impression is good for most of the data volumes, this does not necessarily coincide with very high indices. Unfortunately, the method fails completely for data set number 5 (see the table), where there are almost no visible MS lesions.



(a) (b) © (d) (e)

Figure 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.

Performance indices obtained by leave-one-out validation for all of the examined data sets. From left to right the columns contain the achieved Dice coefficient, Pearson correlation coefficient, sensitivity, specificity, positive predictive value, and negative predictive value.

| | 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 |

## Conclusion

In this paper, a new method for the segmentation of multiple sclerosis lesions in multispectral 3-D brain MRI data was presented. The results show that, even though supervised techniques based on individual voxel intensities for tissue classification within MR data are mostly considered inappropriate due to inter-scan variabilities, this is not necessarily true for structural approaches that additionally consider the context of a voxel for the purpose of classification. 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.

## References


[1] van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A., Suetens, P.: Automated segmentation of multiple sclerosis lesions by model outlier detection. In: IEEE T. Med. Imag. **20**(8) (2001) 677–688

[2] Tu, Z.: Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering. In: IEEE Int. Conf. Comp. 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. Comp. Vis., Graz, Austria. (2006) 436–448

[4] Sonka, M., Fitzpatrick, J.M.: Handbook of Medical Imaging, Vol. 2: Medical Image Processing. SPIE—The Int. Soc. Opt. Eng., 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. In: Ann. Stat. 28 (1998) 337–407

[8] 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

[9] Viola, P., Jones, M.: Robust real-time object detection. In: 2nd Int. Workshop on Stat. Learning and Comput. Theories of Vis., Vancouver, Canada. (2001) 1–25

[10] Besag, J.: On the statistical analysis of dirty pictures. J. R. Stat. Soc. Ser. B Stat. Methodol. **48**(3) (1986) 259–302

Michael Wels was born in 1979 and graduated with degree in computer science from the University of Wuerzburg in 2006. Currently, he is a member of the research staff at the University of Erlangen-Nuremberg's Chair of Pattern Recognition working towards his Ph.D. His research interests are medical imaging in general and the application of machine learning techniques to medical image segmentation.

Martin Huber studied at the University of Karlsruhe and received his Ph.D. degree in computer science in 1999. Since 1996, he has been with Siemens Corporate Technology and Siemens Medical Solutions. He currently is technical coordinator of the EU funded project Health-e-Child with research interests in medical imaging and semantic data integration.

Joachim Hornegger graduated with a degree in computer science (1992) and received his Ph.D. degree in applied computer science (1996) at the University of Erlangen-Nuremberg (Germany). His Ph.D. thesis was on statistical learning, recognition, and pose estimation of 3-D objects. Joachim was a visiting scholar and lecturer at Stanford University (Stanford, CA, USA) during the 1997–1998 academic year, and, in 2007–2008, he was a

visiting professor at Stanford's Radiological Science Laboratory. In 1998, he joined Siemens Medical Solutions Inc., where he was working on 3-D angiography. In parallel with his responsibilities in industry, he was a lecturer at the Universities of Erlangen (1998–1999), Eichstaett-Ingolstadt (2000), and Mannheim (2000–2003). In 2003, Joachim became Professor of Medical Imaging Processing at the University of Erlangen-Nuremberg, and since 2005, he has been chaired professor heading the Chair of Pattern Recognition. His main research topics are currently pattern recognition methods in medicine and sports.