

# High-Precision Computer-Assisted Segmentation of Multispectral MRI Data Sets in Patients with Multiple Sclerosis by a Flexible Machine Learning Image Analysis Approach

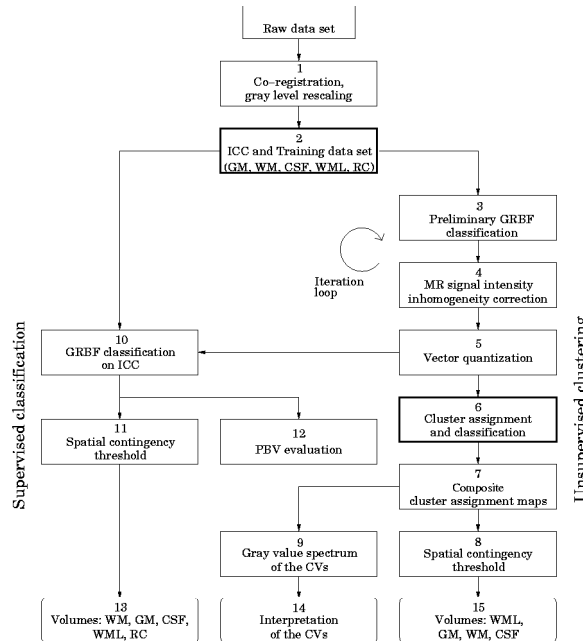
Axel Wismüller<sup>1</sup>, Johannes Behrends<sup>1</sup>, Oliver Lange<sup>1</sup>, Miriana Jukic<sup>1</sup>,  
Klaus Hahn<sup>2</sup>, Maximilian Reiser<sup>1</sup>, and Dorothee Auer<sup>3</sup>

<sup>1</sup>Institut für Klinische Radiologie, Klinikum der Universität München,  
Ludwig-Maximilians-Universität München, Ziemssenstraße 1, 80366 München  
<sup>2</sup>Klinik und Poliklinik für Nuklearmedizin, Klinikum der Universität München,  
Ludwig-Maximilians-Universität München, Ziemssenstraße 1, 80366 München  
<sup>3</sup>Max-Planck-Institut für Psychiatrie, Kraepelinstraße 2-10, 80804 München

**Abstract.** Automatic brain segmentation is an issue of specific clinical relevance in both diagnosis and therapy control of patients with demyelinating diseases such as Multiple Sclerosis (MS). We present a complete system for high-precision computer-assisted image analysis of multispectral MRI data based on a flexible machine learning approach. Careful quality evaluation shows that the system outperforms conventional threshold-based techniques w.r.t. inter-observer agreement levels for the quantification of relevant clinical parameters, such as white matter lesion load and brain parenchyma volume.

## 1 Introduction

In the light of current scientific discussions on the clinical role of MRI for the evaluation of white-matter disease [1], the development of *flexible* innovative strategies for computer-assisted high-precision segmentation methods is a subject of topical interest in human brain imaging. Flexibility here refers to (i) the input, (ii) the output, and (iii) the level of human intervention required in such systems. As far as the input is concerned, the user should have the opportunity of freely choosing among different MRI sequences and various combinations thereof. As for the output, the system should not be restricted to lesion quantification alone, but should offer the potential to provide high-precision whole-brain or tissue-specific segmentation as well, in order to account for global brain atrophy measures, e.g. Percentage of Brain parenchyma Volume (PBV), which have recently moved into the focus of current basic and clinical research interest [2]. Finally, the system should offer different levels of human intervention: On one hand, the development and evaluation of computer-assisted segmentation systems can benefit from the superior image analysis capabilities of human beings which implies a higher degree of operator interaction. On the other hand, for

**Fig. 1.** The segmentation system.

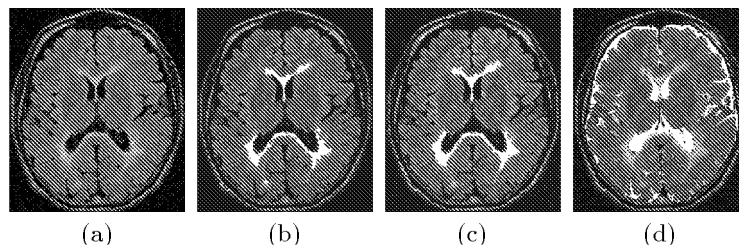
large-scale clinical (e.g. multi-center) studies, however, a reduction of human intervention may sometimes be helpful in situations where user interaction could reduce reproducibility, i.e. could impose subjective bias on segmentation results. Thus, the development, test, and evaluation of a segmentation system aiming at the analysis of specific pathological changes in MS is a challenge that requires considerable effort w.r.t. integrating substantial human expertise in order to optimize computer-assisted decision support. In this paper, we present a neural network-based segmentation system for multispectral MRI data sets of the human brain that has been specifically designed in order to provide a high degree of flexibility with regard to all three aspects mentioned above.

## 2 Methods

**Data:** Six patients with relapsing-remitting MS and EDSS [3] scores between 1.0 and 3.5 were included in the study. Image data were obtained on a 1.5 T MRI scanner **General Electric, Signa™** employing a standardized MRI sequence protocol including  $T1$  and  $T2$  weighted, *Proton Density* (PD) weighted, *Fluid-attenuated Inversion-Recovery* (FLAIR), and *Magnetization Transfer* (MT) sequences in axial slice orientation. The  $T1$  and MT sequences were repeated after intravenous contrast agent administration. Total scanning time was 27.4 min.

**Image Analysis:** The conceptual basis of single components of our system has been described in [4]. Here, we want to put special emphasis on the functional interplay between the various components in so far as it is relevant to

**Fig. 2.** (a) Axial FLAIR slice of a brain containing WML; (b) WML classification based on interactive cluster assignment using the CASCADE system; (c) supervised automatic WML classification using a GRBF neural network; (d) CSF segmentation by GRBF neural network classification. For explanation, see text.



brain segmentation in MS. An overview of the segmentation system is shown in fig. 1. Thick-lined boxes indicate interactive steps. Boxes with rounded corners refer to segmentation results. After co-registration and gray level rescaling (1) of the input data, the intracranial cavity (ICC) is pre-segmented interactively. For the data presented here, this step was performed manually by human expert readers, however, (semi-)automatic techniques may be used as well, such as the methods developed by our group [5]. In a second step, a training data set is obtained manually comprising small reference regions labeled as “Gray Matter (GM)”, “White Matter (WM)”, “Cerebrospinal Fluid (CSF)”, “White Matter Lesion (WML)”, and a “Residual Class (RC)”, representing other tissues such as meninges or larger vessels (2). Subsequently, gray level shift effects induced by magnetic field inhomogeneities and cross-talk effects can be corrected using the training data and the ICC masks (3, 4). For this purpose, we have developed a specific bootstrap algorithm based on iterative improvement of a preliminary neural network tissue classification, which will be published elsewhere. After these preprocessing steps, each voxel within the ICC mask is assigned to a feature vector  $\mathbf{x}$  representing its MRI signal intensity spectrum. This set of feature vectors is partitioned into  $N$  clusters by unsupervised learning (5) based on minimal free vector quantization [4]. The resulting codebook can either be used for interactive visual tissue type classification based on cluster assignment maps (6), or automatic supervised segmentation can be obtained by subsequent training of a Generalized Radial-Basis Functions- (GRBF-) neural network (10), see [4].

For the interactive visual classification of cluster assignment maps, we developed a software system named CASCADE (Computer-Assisted Cluster Assignment Decision Environment) which enables quick and efficient screening of cluster assignment maps and underlying MRI data. Here, each feature vector  $\mathbf{x}$  is uniquely attributed to its closest codebook vector  $\mathbf{w}_j(\mathbf{x})$  according to a minimal distance criterion in the gray level feature space, and corresponding cluster assignment maps (6) are constructed for visual inspection. In a second step, each cluster  $j$  belonging to codebook vector  $\mathbf{w}_j$  is interactively assigned to a specific tissue class  $\lambda \in \{0, \dots, m\}$  by a human expert reader. Finally, all the clusters assigned to each specific tissue class  $\lambda$  are collected and merged yielding a composite cluster assignment map (7) representing the final segmentation result. Based

**Table 1.** Statistical analysis of WML and PBV quantification methods w.r.t. inter-observer agreement (univariate  $F$ -test,  $N = 6$ ). The method yielding better results, i.e. higher inter-observer agreement is printed in bold face for each pairwise comparison.

Method A	Method B	$p$ -value
WML segmentation		
Region Growing	<b>GRBF</b>	0.075
Region Growing	<b>CASCADE</b>	0.029
<b>GRBF</b>	<b>CASCADE</b>	0.027
PBV calculation		
<b>GRBF</b>	Angle Image	0.003

on this tissue assignment, the isolated or merged codebook vectors representing prototypical gray level spectra may be plotted for further visual analysis and interpretation (9, 14). Finally, segmentation results may be used for tissue-specific volume measurements (15), where spatial smoothing techniques or geometric contingency thresholding (8) can be employed as optional post-processing steps. An example for WML segmentation results is presented in fig. 2b.

Alternatively, for automatic supervised classification by a GRBF neural network (10) the training data from step (2) and the resulting codebook from step (5) can be re-cycled [4]. Based on the respective tissue segmentation, the WML volume can be quantified as well (11, 13 – see fig. 2c). Furthermore, the GRBF segmentation approach can be used for PBV calculation based on automatic CSF identification (12 – see fig. 2d).

### 3 Evaluation and Results

In order to perform a thorough quantitative evaluation of the described preprocessing and segmentation procedures w.r.t. all data processing steps involving human interaction, WML quantification and PBV computation were performed based on (i) interactive definition of training data sets by two different observers independently for supervised GRBF classification of WML and PBV computation, respectively, (ii) interactive reference region contour tracing for threshold definition of an observer-guided region growing technique [7] serving as a reference method, by two different observers independently, (iii) interactive cluster assignment using the CASCADE system by two different observers independently, and (iv) interactive threshold definition for the angle image method [2] serving as a reference method for PBV computation, by two different observers independently. The computation of inter-observer agreement levels was performed according to the statistical guidelines of the British Standards Institution [6]. In order to rank the methods w.r.t. their segmentation quality, the inter-observer agreements of CASCADE and GRBF neural network segmentation were compared to region growing, based on a univariate  $F$ -test. From the results presented in tab. 1, it can be concluded that (i) the mean inter-observer agreement in cluster assignment using the CASCADE segmentation procedure is higher than in both region growing ( $p = 0.029$ ) and GRBF neural network classification ( $p = 0.027$ ), i.e. there is a significant method effect; (ii) the mean

inter-observer agreement in GRBF neural network classification is higher than in threshold-based region growing. However, statistical analysis reveals only a method effect of reduced significance for the comparison of GRBF neural network segmentation and region growing ( $p = 0.075$ ). In conclusion, interactive cluster assignment using the CASCADE segmentation system performs significantly best in a comparison of the three methods, whereas supervised GRBF neural network classification is slightly better than conventional region growing serving as a reference method for WML quantification. For PBV computation, our GRBF neural network method outperforms the reference angle image technique w.r.t. inter-observer agreement at a significance level of  $p = 0.003$ .

## 4 Discussion

For WML quantification in MS we obtain the best segmentation results using the CASCADE approach, where human expert knowledge is incorporated at a “cluster level” instead of a “pixel level”, i.e. at an advanced, abstract level of knowledge representation within the pattern recognition process. We conjecture that this observation could be of particular interest in the light of ongoing discussions on “domain knowledge data fusion for decision support” in the machine learning community. Our study shows that computer-assisted image analysis using semi-automatic neural network segmentation outperforms conventional threshold-based techniques w.r.t. inter-observer agreement levels for both WML quantification and PBV calculation in MRI data of MS patients. At the same time, our segmentation system allows the radiologist and neuroscientist to choose freely among different input MRI sequences and various combinations thereof in order to systematically explore their contribution to brain imaging in MS.

## References

1. Filippi M, Horsfield M, Tofts P, et al.: Quantitative assessment of MRI lesion load in monitoring the evolution of multiple sclerosis. *Brain*, 118:1601–1612, 1995.
2. Ge Y, Grossman RI, Udupa JK, et al.: Brain atrophy in relapsing-remitting multiple sclerosis and secondary progressive multiple sclerosis: longitudinal quantitative analysis. *Radiology*, 214:665–670, 2000.
3. Kurtzke JF: Rating neurologic impairment in multiple sclerosis: an expanded disability status scale EDSS. *Neurology* 33:300–311, 1983.
4. Wismüller A, Vietze F, Dersch DR: Segmentation with neural networks. In: Bankman I, et al. (eds.), *Handbook of Medical Imaging*, Academic Press, ISBN 0-12-077790-8, 2000.
5. Wismüller A, Behrends J, Lange O, et al.: Automatic segmentation of cerebral contours in multispectral MRI data sets of the human brain by self-organizing neural networks. *Radiology* 221(P):461, 2001.
6. British Standards Institution: Precision of test methods. Part I. Guide for the determination of “reproducibility” for a standard test method by inter-laboratory tests. BS 5497. 1987.
7. Wicks D, Tofts P, et al.: Volume measurement of multiple sclerosis lesions with magnetic resonance images: a preliminary study. *Neuroradiology* 34:475–479, 1992.