Recognition of grey matter and parallel versus crossing fibre bundles within white matter using HARDI data and a supervised learning algorithm.

S. SCHNELL[1], B. W. KREHER[1], J. HENNIG[1], H. BURKHARDT[2], AND V. KISELEV[1]

[1]DEPT. OF DIAGNOSTIC RADIOLOGY, MEDICAL PHYSICS, UNIVERSITY HOSPITAL FREIBURG, FREIBURG, GERMANY, [2]DEPT. OF COMPUTER SCIENCE, CHAIR IN PATTERN RECOGNITION AND IMAGE PROCESSING, ALBERT-LUDWIGS UNIVERSITY FREIBURG, FREIBURG, GERMANY

INTRODUCTION: Diffusion weighted MRI provides biologically relevant information about the tissue microstructure. Information about microstructural parameters is derived from biophysical modelling of the signal formation in investigated tissues. However, for a realistic tissue description such models are rather complex and often include ill defined parameters not adequately supported by the measurement data. As an alternative approach data-driven analysis of multidirectional diffusion weighted MRI data may provide unique fingerprints for different types of tissue. Such an analysis can be performed using efficient algorithms for pattern recognition known from computer science. In the present study we use a support vector machine (SVM) [1], a supervised machine learning technique for classification, which has demonstrated a robust performance in a number of applications [e.g. 2]. The SVM is used for differentiation of three different classes: grey matter (GM), fibre crossing bundles (CR) and parallel fibre bundles (PF) within white matter (WM). Measurement data are gathered from high angular resolution diffusion imaging (HARDI). The analysis requires: first, an appropriate representation of the data; second, the creation of labelled data sets for training (containing pairs of input data and desired outputs) and testing; and third, the application to *in vivo* data. HARDI data are represented by rotational invariant weights of low-order spherical harmonics.

The method was systematically tested using simulations and applied to an *in vivo* data set. Verification of assignment to GM was based on a T1-weighted (MPrage) dataset. *In vivo* validation of PF versus CR would require a gold standard, which is not available. Therefore a simulated data set was used for training of the algorithm.

METHODS: The diffusion weighted images were simulated for 81 diffusion encoding directions (DEdirs) evenly distributed on a sphere. The same template was used for the simulation of GM, PF and CR, which included two crossing fibre bundles with varying angle between them and relative weights in the signal. The fractional anisotropy (FA) and the mean diffusivity (mD) were equal for both bundles. The mean diffusivity was constant for each tissue type: $0.39*10^{-3}$ mm²s^{-1} in GM and $0.79*10^{-3}$ mm²s^{-1} in PF and CR. The remaining three parameters were varied independently within each class. GM was simulated with small FA up to 0.15, the relative weight and the angle between fibres were close to zero. For PF a FA above 0.55 was assigned with an angle between fibre bundles close to zero. The simulation of CR was performed with an intermediate FA from 0.2 to 0.55, a relative weight of fibres of 0.2 to 0.5 (both bundles are the same) and with an increasing angle above 30°. Six data sets with different noise instants (Gaussian white noise) were created as training data sets. The testing data sets were identical except that the fibres were randomly rotated about the z-axis and that new noise instants were added. 520 parameter combinations representing one tissue class were simulated.

A prerequisite of the tissue recognition using the SVM is a rotational invariant data representation. The information about fibre orientation was removed from by expansion in spherical harmonics (SH) and computing the power of the orders zero to six. A SVM C++ library called libSVMtl [3] was used, which has a number of internal algorithmic options.

For a test of performance of the SVM using *in vivo* data, HARDI data was acquired on a 3 T Siemens TRIO (TR = 11000 ms, TE = 94 ms, voxel: 2x2x2 mm³, matrix: 104x104 pixel², DEdirs = 81 with b = 1000, 11 b = 0 images). An additionally acquired T1-weighted MPrage was coregistered to the non-weighted HARDI images and segmented into GM, WM and cerebrospinal fluid. The SVM was trained with a simulated data set (SNR = 10).

RESULTS: Fourteen different SVM specific algorithmic options were tested. Amongst those, four algorithms all using a radial basis function as kernel showed the best sensitivity and specificity. Only the results of the two best algorithms are presented in table 1. The SVM demonstrated an excellent performance. Most critical was the detection of PFs. For algorithm 1 the PFs were wrongly detected to be crossings at an SNR level of 100 in a single parameter set. The detection of CR showed small error quota starting at an SNR of 30. At the lowest SNR level of five GM detection remained still above 99 %. The classification resulting from algorithm 2 was similar to algorithm 1, but had a higher error quota for PF in the lower SNR range.

	GM						PF						CR					
SNR	∞	100	60	30	10	5	∞	100	60	30	10	5	∞	100	60	30	10	5
algorithm 1:																		
true positives	100,00%	100,00%	100,00%	100,00%	100,00%	99,42%	100,00%	99,81%	99,81%	99,81%	97,88%	92,69%	100,00%	100,00%	100,00%	98,85%	99,23%	95,96%
false positives	0,00%	0,00%	0,00%	0,00%	0,00%	0,48%	0,00%	0,00%	0,00%	0,58%	0,38%	1,63%	0,00%	0,10%	0,10%	0,10%	1,06%	3,85%
algorithm 2:																		
true positives	100,00%	100,00%	100,00%	100,00%	100,00%	99,62%	100,00%	100,00%	100,00%	99,81%	95,96%	89,62%	100,00%	100,00%	100,00%	99,81%	99,81%	97,69%
false positives	0,00%	0,00%	0,00%	0,00%	0,00%	0,38%	0,00%	0,00%	0,00%	0,10%	0,10%	0,87%	0,00%	0,00%	0,00%	0,10%	2,02%	5,29%

Table 1: Results of segmentation in simulated phantoms with six different noise instants. The results of the two best performing SVM algorithms are listed.

The results for the segmentation of the *in vivo* data are shown in the figures below. There is no quantitative evaluation of correctness possible; one expects that about 90 % of the WM are crossings. Figure 1a highlights crossings (CR) and 1b the parallel fibre bundles (PF) both overlaid on FA maps. The results were masked with WM masks obtained from the T1-weighted image. Figures 1c and d show masks created by thresholding FA.

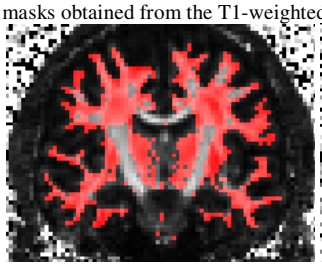

Figure 1a: Found CR in *in vivo* data illustrated in red.

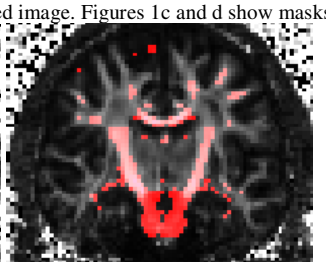

Figure 1b: Found PF in *in vivo* data illustrated in red.

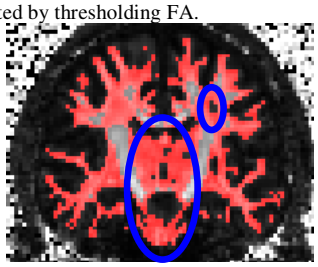

Figure 1c: FA mask thresholded between 0.2 - 0.55. The blue ellipses indicate errors, cf. fig. 1a.

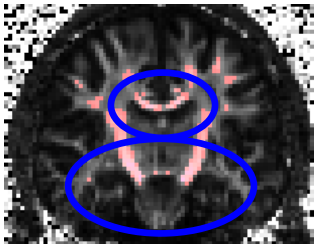

Figure 1d: FA mask thresholded above 0.55. The blue ellipses indicate errors, cf. fig. 1b.

DISCUSSION AND OUTLOOK: A new method for the separation of parallel and crossing fibre bundles in the brain white matter using HARDI data and a SVM algorithm is presented. A rotationally invariant data representation was used as input for the SVM. The chosen SVM algorithms proved a surprisingly high accuracy and robustness even with data sets having a SNR of only five. The recognition of fibre crossings is based on HARDI data obtained in a typical clinical setting. T1 weighted images were used to validate the recognition of grey matter. Our analysis showed to be more sensitive than FA thresholding especially for crossings (cf. figs 1a - d), which is due to the usage of SH orders above two giving more information about the underlying diffusion distribution. The FA value relies on the diffusion tensor model, which correlates with the SH of order two. Proper verification of the results remains a challenge, which is a commonly recognised problem in DTI fibre tracking. Our method can provide *a priori* knowledge for increasing the performance of fibre tracking algorithms. Future work will be dedicated to derive further refinements for sub-classification of CR and GM areas. Also, SVM allow including more and different types of MRI data for multi-parametric tissue classification.

REFERENCES

[1] C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, Kluwer Academic Publishers, Boston, Netherlands, 1998

[2] LG Durand et al., Comparison of pattern-recognition methods for computer-assisted classification of spectra of heart sounds in patients with a porcine bioprosthetic valve implanted in the mitral position, IEEE Transactions on Biomedical Engineering 37 (12): 1121-1129 DEC 1990

[3] O. Ronneberger, University Freiburg, Institute of Computer Science, 2004, http://lmb.informatik.uni-freiburg.de/lmbsoft/libsvmtl/

ACKNOWLEDGEMENTS: The work was supported by the Bundesministerium für Bildung und Forschung [BMBF-research collaborations "Mechanisms of brain reorganisation in the language network" (01GW0661)].