

Machine-based Rejection of Low-quality Spectra and Estimation of Brain Tumor Probabilities from Magnetic Resonance Spectroscopic Images

Bjoern H. Menze¹, B. Michael Kelm¹,
Daniel Heck¹, Matthias P. Lichy^{2,3} and Fred A. Hamprecht¹

¹Interdisziplinäres Zentrum für Wissenschaftliches Rechnen (IWR),
Universität Heidelberg, 69120 Heidelberg

²Abteilung Radiologie, Deutsches Krebsforschungszentrum (dkfz), 69120 Heidelberg

³currently: Diagnostische Radiologie, Universität Tübingen, 72076 Tübingen

Abstract. Magnetic resonance spectroscopic images (MRSI) carry spatially resolved information about the *in vivo* metabolism, however, their evaluation is difficult. Problems arise especially from artifacts and noise, yielding non-evaluable signals in many voxels. We propose a two-step approach to the processing of MRSI. In the first step a non-linear classifier is employed in every voxel to determine whether the spectral signal is evaluable, and if so, the tumor probability is computed in the second step. Thus, the quality control is strictly separated from the diagnostic evaluation of the spectrum. For an assessment of the proposed approach we consider MRSI-based brain tumor detection and localization and a tumor probability mapping by pattern recognition. In a quantitative comparison against the standard operator-controlled processing our interaction-free approach shows similar to superior performance.

1 Introduction

Magnetic resonance spectroscopy (MRS) allows the study of physiological and patho-physiological processes in the living tissue. Thus, characteristic changes in the relative concentration of a number of metabolites can be detected by MRS. Magnetic resonance spectroscopic imaging (MRSI) also allows for a spatial mapping of their occurrence. – One of the major applications of MRSI in clinical routine is in the diagnosis of tumor, in particular in the brain, but also in the breast or prostate.

MRSI in tumor diagnosis

The standard approach in the analysis of MR spectra is a fitting of resonance-line shaped template functions to the spectral pattern. The parameters of these (amplitude, line width, area) provide the basis for the following diagnostic interpretation of the spectrum, e.g. using the value of the resonance line integral of Choline relative to the integral of Creatine as an indicator for the malignity

of a tissue change. Unfortunately, the robustness of this resonance line “quantitation” is limited in the presence of noise or artifacts within the spectra, and a visual control by the operator is always required. This effort hampers the routine use of MRSI, but is compulsory to assess the diagnostic value of an MR spectrum.

An alternative approach, which is primarily suited for highly resolved (short echo time) spectra, bypasses the quantitation and uses algorithms from pattern recognition and machine learning. Tissue type and state of disease are directly inferred on the spectral pattern, without a model based reduction of the spectral information. Pattern recognition on the full spectrum, e.g. by a projection of the spectral image to a “nosologic” or diagnostic map, is more robust than the fitting of adaptive resonance line functions [1]. Nevertheless, this method also fails in the presence of spectral artifacts or at low signal-to-noise ratio. So, an inspection of the spectra is also required here and is limiting a fully automated analysis also for this approach.

Published examples for the use of MRSI in tumor diagnostic can be found for both analysis approaches, e.g. in the non-linear classification of brain tumor [2] (EU-project INTERPRET), or in the quantitation of MR spectroscopic images in the detection of recurrent tumors after radiation therapy [3]. The most important obstacle for an application of these automated routines is the lack of a specific assessment about the confidence in the diagnostic analysis of a spectrum. While a number of algorithms are available for the diagnostic interpretation of MR spectra [2, 3, 4], quality control and the evaluation of confidences still demand a large amount of operator time and bind the resources of the human expert.

Two step analysis

Using the example of the detection and localization of brain tumor, we propose in the following a two-step procedure which is able to overcome some of the mentioned limitations in the automated analysis of MR spectra. In the first step, we demonstrate the applicability of pattern recognition in the classification of evaluable and non-evaluable (“nice vs. noise”) spectra. In the following diagnostic analysis, we use regression methods from chemometrics for pattern recognition and the generation of probabilistic tumor-maps with confidence assessment.

Overall, we present an automated processing of MR spectroscopic images which yields a data product similar to any other standard imaging modality. Most important in comparison to the standard processing of MRSI, it is not a collection of localized spectra that need to be interpreted laboriously, but the mapping of the tumor-related information within the spectral image.

Results of this approach will be compared against the outcome of a operator-controlled analysis. In order to allow a direct application, the algorithms can be integrated into the radiation therapy planning software at the German Cancer Research Center (dkfz) [5].

2 Methods

Data

The data was acquired on a standard MR scanner (Siemens, Magnetom Vision, 1.5T, ^1H -MRS, TE 135 ms) at the dkfz. Spectral images of 31 patients with tumor diagnosis (astrocytoma, meningioma, glioma, metastases – confirmed during follow-up) were available at a spatial resolution of 32^2 and under a spatial resolution of approx. 1 cm voxel side-length (fig. 1a). Within this data 269 spectra could be labeled as “tumor”, “tumor border” or “normal state” (69/49/151, fig. 1d). – Spectral images of another 36 patients were classified completely by visual inspection according to the signal quality and resulted in approx. 37000 spectra (10% “nice”) for the training of the “nice vs. noise” classifier.

Classifier – quality assessment

While the signal-to-noise ratio can be assessed easily by linear methods, the separation of the various artifacts demands for the training of a non-linear classifier (“randomForest” decision tree ensemble [1]). RandomForests extend the idea of bagging tree ensembles by seeking optimal splits for each tree node on random subspaces of the data. Outcome of a randomForest classification is either the binary class membership, or the normalized number of votes from the tree ensemble, which allows for a subjective readjustment of the rejection threshold. – Different feature representations of the data (different normalization strategies, data binning) were tested and optimized according to the classification error. As the “noise” class of the training data only contained the most frequent artifacts, it was extended by uniformly distributed random samples of the spanned feature space of all spectra. This nonparametric outlier detection ensures a proper signal classification even in the presence of previously unknown artifacts.

Classifier – tumor detection

The character of the binary “normal state vs. tumor” classification problem allows the application of linear models on the data. In order to overcome the sparseness of the training problem (256 spectral channels (P) in the spectral region of Choline, Creatine and NAA, 269 (N) highly collinear spectra), the spectral regions relevant for the classification were identified in a first feature selection by help of the randomForest Gini importance measure (fig. 1d), the accumulated Gini decrease for all nodes of all trees within the ensemble. A partial least squares regression (PLS) [1] on these variables was used for the discrimination of the two tissue types. To allow an assessment of the tumor probability, a logistic model was build on the score of the PLS subspace (fig. 1c).

The leave-‘one-patient’-out cross-validated PLS score was compared by means of the receiver-operator-characteristic (ROC) against the results of different standard approaches [5], all based on resonance line quantitation (line integral ratios and linear combinations thereof, fig. 1b).

3 Results

Areas with high signal-to-noise ratio were reliably identified, virtually all artifacts being present in the data set were separated. An inspection of the leave-one patient-out shows that the cross-validated test error of 1.04 % is most probably due to ambivalent training labels. The classifier proposed to reject 60 low quality spectra (16/10/34) from the data set for the tumor classification.

Within that data set, the randomForest Gini importance identified the spectral regions of the metabolites choline and NAA to be relevant for the classification. Thus a subset of $P_{subset} = 14$ out of the 256 spectral channels was used for training and evaluation of the binomial PLS model (fig. 1d). The resulting tumor classifier reached an accuracy of 95.5% in the differentiation between normal state and both tumor groups, normal and tumor were separated without error (100% accuracy). These values, however, reduce to 92.5% and 93.3% when the 60 low quality spectra are included.

Comparing the ROC of this approach with other methods as described in [5], we find that the automated pattern recognition performs comparable to the outcome of a quantitation by a human operator (fig. 1b with noise) or better (fig. 1b without noise). All studied methods tend to show a higher reliability when low quality spectra are separated by the first step classifier.

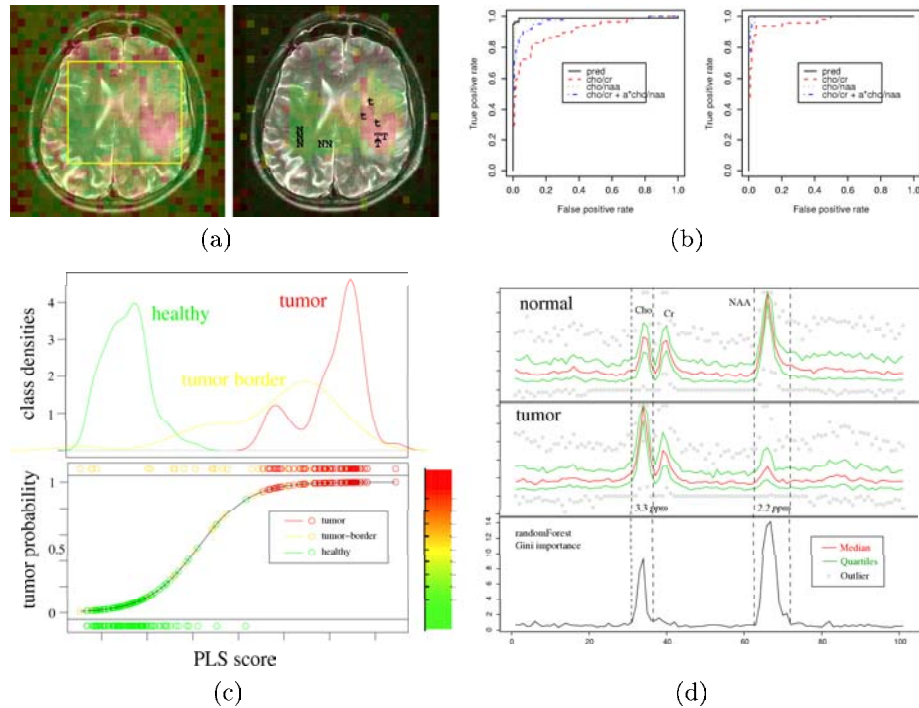
4 Discussion

The proposed application of a specific, non-linear classifier for the the separation of non-reliable spectra is independent from the diagnostic question and the examined organ. Although we observe optimal results in conjunction with our pattern recognition approach, such a quality assessment prior to the diagnostic analysis shows to be beneficial also in conjunction with traditional resonance line quantitation. So far, only information theoretic criteria (e.g. Cramer-Rao-bounds) were used to assess the quality of a spectrum. Explicit and easily available knowledge about frequent artifacts which is the base of our noise classifier are disregarded by these traditional quality measures.

In the specific task of brain tumor detection and localization, we can show that the fully automated pattern recognition operates as reliably as the operator-supervised resonance line quantitation. The high degree of automation of both steps, the trained quality control and the chemometrical tumor recognition, allows for the application of further image processing and the exploitation of spatial relations within the MR spectroscopic image. Promising extensions might be the automated segmentation of tumor regions or the combination of MRSI nosologic images with other diagnostic imaging modalities.

The operator-free data processing is a prerequisite for the clinical application of spectral images with high spatial resolution and therefore thousands of spectra in each acquisition.

Fig. 1. **a)** Tumor probability map; left: PRESS box indicated, right: gold standard (T (t) - tumor (-border), N - normal), regions which were classified as low quality spectra were disregarded for display. **b)** Receiver-operator-characteristic of the presented PLS based pattern recognition and other methods as described in [5], including (left) and excluding (right) low quality spectra. **c)** Transformation of class densities and binomial tumor-probability in red-green color scale. (here: normal vs. both tumor groups) **d)** Average spectral patterns of healthy and tumorous tissue.



References

1. Menze BH, Lichy MP, Bachert P, et al. Optimal classification of Long Echo Time In Vivo Magnetic Resonance Spectra in the Detection of Recurrent Brain Tumors. NMR in Biomedicine submitted;
2. Devos A, Lukas L, Suykens JA, et al. Classification of brain tumor using short echo time ^1H MR spectra. J Magn Reson 2004;p. 164–175.
3. Lichy MP, Pathow C, Schultz-Ertner D, et al. Follow-up of gliomas after radiotherapy: ^1H MR spectroscopic imaging for increasing diagnostic accuracy. Neuroradiology 2005;p. 1–12.
4. Kelm BM, Menze BH, Neff T, et al. CLARET: a tool for fully automated evaluation of MRSI with pattern recognition methods. In: Procs BVM; 2006. .
5. Neff T, Bachert P, Lichy MP, et al. Einsatz der Protonten-MR-spektroskopischen Bildgebung zur Zielvolumenbestimmung in der Strahlentherapieplanung bei glialen Hirntumoren. In: Procs BVM; 2004. p. 145–149.