Biomed Tech 2010; 55 (Suppl. 1) © 2010 by Walter de Gruyter • Berlin • New York. DOI 10.1515/BMT.2010.308

# Brain Tissue Classification Based on Spectroscopy Data

Benjamin Roeschies[1], Martin Busch [2], Susanne Winter[1]
[1]Ruhr-Universität Bochum, Institut für Neuroinformatik, Germany
[2]Grönemeyer Institut für Mikrotherapie, Bochum, Germany

Benjamin.Roeschies@neuroinformatik.ruhr-uni-bochum.de

## Kurzfassung

In dieser Arbeit stellen wir Methoden zur Klassifikation von Gehirngewebe als tumorös oder gesund vor. Ein besonderes Augenmerk liegt dabei auf der Unterscheidung zwischen Astrozytomen und Glioblastomen im Falle tumorösen Gewebes. Diese Unterscheidung ist in sofern bedeutend, als dass sie große Auswirkungen auf die Prognose und die spätere Therapie des Patienten hat. Wir haben drei verschiedene Ansätze zur Lösung dieser Aufgabe verglichen, im einzelnen sind das die lineare Diskriminanzanalyse (LDA), Supportvektormaschinen (SVM) und künstliche neuronale Netwerke (ANN). Die Klassifikation wurde basierend auf den relativen Konzentrationen der wichtigsten Metaboliten des Gehirns im betreffenden Gewebe durchgeführt, welche mit Hilfe einer Protonenresonanzspektroskopie mit kurzer Echozeit bestimmt wurden. Wir konnten sehr zufriedenstellende Resultate erzielen mit nahezu 100% korrekten Entscheidungen zwischen tumorösem und gesundem Gewebe und etwa 90% korrekten Entscheidungen zwischen Astrozytomen und Glioblastomen.

## Abstract

We present methods for classification of brain tissue as tumorous or healthy in this paper with an additional emphasis on distinguishing between astrocytoma and glioblastoma in case of tumorous tissue, which is an important decision as it has large impact on the following therapy and prognosis for the patient. We compared three different approaches to this task, namely linear discriminant analysis (LDA), support vector machines (SVM), and artificial neural networks (ANN). The classification was performed based on the relative concentrations of the most important metabolites within the respective tissue. These concentrations were gathered using proton magnetic resonance spectroscopy with short echo times. We received very satisfying results with almost 100% correct decisions between tumorous and healthy tissue and about 90% correct decisions between astrocytoma and glioblastoma.

## 1 Introduction

The most reliable method for brain tissue classification is currently a biopsy, which has the major drawbacks of being invasive and may have serious consequences for the patient. In contrast to that, proton magnetic resonance spectroscopy is non-invasive and can be performed with much less effort [2, 6]. Therefore, it is desirable to use the data provided by this method for a reliable tissue classification. It has already been shown that the distinction between tumorous and healthy tissue is possible [5]. Fig. 1 shows an example of the data acquired by proton magnetic resonance spectroscopy.

In doing the classification it is not only important to decide between tumorous and healthy tissue, but also to differentiate tumor types and grades as this has a crucial impact on prognosis and therapy for the patient. The latter decision turns out to be a lot more difficult.

In this work we used three different classification methods to solve this task: support vector machines (SVM), artificial neural networks (ANN), and linear discriminant analysis (LDA). While support vector machines represent a highly sophisticated mathematical classification method, neural networks are biologically inspired and imitate the behaviour of brain cells. Linear discriminant analysis is a

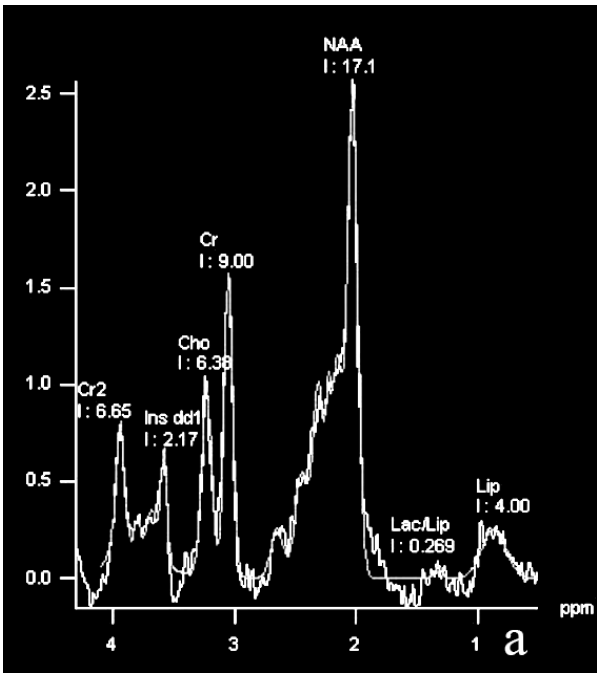


**Fig. 1** Example of spectroscopy data (healthy tissue). The value of I is proportional to the relative concentration of the respective metabolite.

very basic linear classification method that can be seen as a baseline for performance measures.

# 2 Materials and Methods

## 2.1 Data and Experimental Setup

For our experiments we used a total of 350 spectra obtained by proton magnetic resonance spectroscopy of which 231 were labelled as healthy tissue, 30 as astrocytoma, and 89 as glioblastoma. For the two tumor types the categorization was supported by a biopsy. The data was divided into a training set and a test set. We performed tests using different sizes for the training set reaching from as few as 6 samples up to more than 290 samples. The remaining samples were always used as the test set. For each classification method and each training set size classification errors were averaged over 1000 trials to learn the classification.

## 2.2 Linear Discriminant Analysis

The linear discriminant analysis is a method from multivariate statistics that separates data linearly in the input space [3]. The basic principle is to assume a normal distribution of features for each class and estimate a covariance matrix of the features. The separable is then given by Fisher's linear discriminant, which has the property to maximize variance between the two classes and to minimize variance within each class.

## 2.3 Support Vector Machines

Support vector machines are kernel methods that do a linear classification in a so-called feature space [7]. While the classification itself is a simple linear separation, the more important part of this algorithm is the transformation of the input into the feature space, which is done using the kernel function. Support vector machines are maximum margin classifiers, which means the separating hyperplane in the feature space has maximum distance to the images of the training data points in that space.

## 2.4 Artificial Neural Networks

Neural networks are widely used for pattern recognition tasks [1]. We used a special type of neural network for our experiments also known as the multi-layer perceptron (MLP). The neural network consists of one input node for each feature component. Additionally, it uses 7 neurons in a hidden layer. Classification results are represented in terms of vectors, which means there are two output neurons – one for each class – and the neuron with activation closer to 1 decides which class is represented.

The network is adapted to the training data using the improved Resilient-backpropagation-algorithm with weight-backtracking. This is an error gradient decent algorithm with adaptive step size and the possibility to reset network weights to previous states if the output error grows from one time step to the next. The adaptive step size causes the decent to slow down close to an error minimum.

# 3 Results

## 3.1 Tumor Classification

All three classification approaches performed very well concerning the decision between tumorous and healthy tissue. The support vector machines and artificial neural networks reached a median of 98.3% correct classifications for this task using the maximum number of training samples with an inter quartile range (IQR) of 1.7% for both algorithms. Even the very basic linear discriminant analysis came close to those results reaching 96.6% correct classifications with an IQR of 3.4%. The results were probed for significance using a one-sided Wilcoxon rank sum test. While the difference between LDA and the more sophisticated approaches was statistically significant, the absolute difference was rather small. All three methods managed to achieve 100% correct classifications in many of the training trials.

Support vector machines and neural networks had a median specifity of 100% (IQR 2.6%), which means that in the vast majority of the training trials the system recognized healthy tissue with 100% accuracy, while LDA only achieved 97.4% (IQR 2.6%). Sensitivity was 95% for all three approaches on this task, which means that the worse overall performance of LDA is a result of its worse specifity. LDA even achieved a lower IQR with regards to sensitivity with 5% compared to 10% for the other two algorithms.

Fig. 2 shows the accuracy of the 3 approaches for different sizes of the training set. Even with only 6 training samples SVM and neural networks reached accuracies of about 88% on average. At only 17 training samples all three approaches achieve more than 90% accuracy.

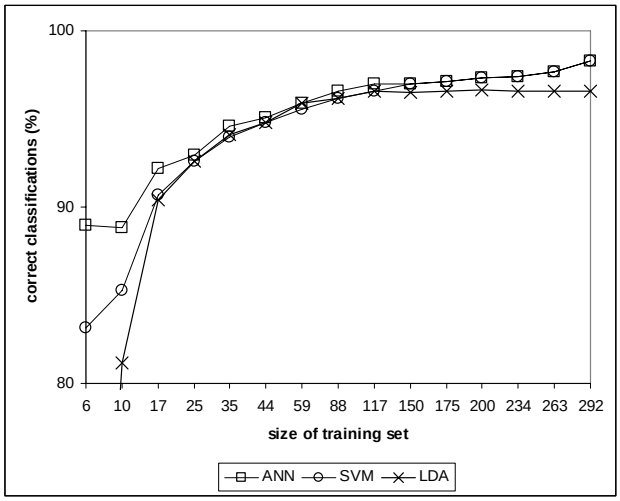


**Fig. 2** Accuracy of the three approaches for different sizes of the training data set at the tumor classification task.

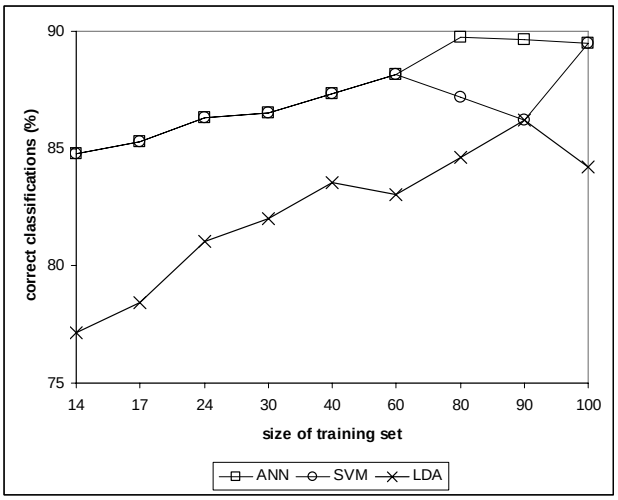


**Fig. 3** Accuracy of the three approaches for different sizes of the training data set at the tumor type classification task.

### 3.2 Tumor Type Classification

The distinction between astrocytoma and glioblastoma in case of tumorous tissue showed slightly worse results and larger differences between the classification methods. SVM and neural networks achieved a median of 89.5% correct classifications on this task while LDA classified correctly in only 84.2% of the cases using the maximum number of training samples. The IQR was 5.3% for the SVM and 10.5% for the neural network and LDA. As in the previous experiment the differences of the medians turned out to be significant.

In contrast to the first task LDA achieved very good specifity with a median of 100% but had its weakness in sensitivity, which was only 85.7% (IQR 7.1%). The more sophisticated approaches were much better at recognizing the glioblastoma reaching 92.9% sensitivity, but often failed in recognizing astrocytoma reaching only 80% as a median. Concerning sensitivity the IQR was 7.1% for the SVM and 14.3% for the neural network. The IQR was much higher concerning specifity with 40% for the SVM and neural networks and 20% for LDA.

Fig. 3 shows the accuracy of the 3 approaches for different sizes of the training set. SVM and neural networks achieve more than 85% accuracy using only 17 training samples.

## 4 Discussion

We have shown that proton magnetic resonance spectroscopy provides a very good indicator for classification of healthy and tumorous tissue. Also, the distinction between astrocytoma and glioblastoma can still be made with good accuracy. In the future this will hopefully lead to a reduction in the need for biopsies in brain tumor diagnosis.

The classification results showed that the spectroscopy provides very good classification features because the difference between the very basic LDA algorithm and the more sophisticated methods turned out to be quite small. However, for the more difficult problem of deciding between tumor types this difference enlarges. Additionally, support vector machines and neural networks can learn successfully from less training samples, which is an advantage especially for medical problems where a large amount of training data is often not available.

## 5 References

[1] Bishop CM: Neural Networks for Pattern Recognition. Oxford University Press, 1995

[2] Devos A et al: Classification of brain tumours using short echo time 1H MR spectra. J Magn Reson., Vol. 170, No. 1, pp. 164-175, 2004

[3] Hastie T et al: The Elements of Statistical Learning. Data Mining, Inference, and Prediction. Springer-Verlag, 2001

[4] Igel C and Hüsken M: Empirical Evaluation of the Improved Rprop Learning Algorithm. Neurocomputing, Vol. 50, No. C, pp. 105-123, 2003

[5] Liebenrodt K et al: Protonenresonanzsprektroskopie des Gehirns mit kurzer Echozeit: Unterstützung der Gewebeklassifizierung durch künstliche neuronale Netzwerke. Biomedizinische Technik (BMT), Vol. 52 (Suppl.), 2007

[6] Tate AR et al: Automated classification of short echo time in vivo 1H brain tumor spectra: a multicenter study. Magn Reson Med. Vo. 49, No. 1, pp. 29-36, 2003

[7] Vapnik V and Chervonenkis A: Theory of Pattern Recognition. Nauka, 1974