A multi-stage classifier for *in vivo* magnetic resonance spectra of brain tumours

J. Minguillón, A.R. Tate, J. Griffiths, C. Arús

April 21, 2002

INTRODUCTION

We are developing a decision support tool, based on a large "training" database of spectra to help clinicians use MRS to categorise brain tumours (`http://carbon.uab.es/INTERPRET`). Automated pattern recognition techniques allow new cases to be displayed and compared with those already in the database.

Previous approaches, classifying only the most common types of brain tumour, are not practical for this application, where all types of brain tumour will be displayed. We present a more usable approach using a tree-based multistage classifier.

MATERIALS AND METHODS

1.5T short-echo (20-32 ms) spectra from three centres St George's Hospital Medical School, London IDI, Bellvitge, Barcelona; Centre Diagostic Pedralbes, Barcelona. 189 Spectra from 81 astrocytomas (18 grade II (ASTII), 6 grade III (ASTII), and 57 grade IV (GLB)), 32 metastases (MET), 37 meningiomas (MN), 6 oligodendrogliomas (OD), 6 lymphomas (LYM), 5 Primitive neuroectodermal tumours (PNET), 4 schwannomas (SCH), 4 haemangioblastomas (HB) and 14 normal volunteers. Protocols and processing are described in [1]. Figure 1 shows a typical ASTII spectrum.

Each stage uses a particular set of classification features, selected using a combination of statistical analysis, splitting performance and previous knowledge. Two classifiers are built: a k-NN classifier (KNN) and a decision tree (DT) and results are combined using a voting scheme (V): if both classifiers agree, generate such label, otherwise, generate "unknown". When enough data is available, LDA is used as the first stage of a cascading ensemble with DT. Figure 2 shows the combined scheme.

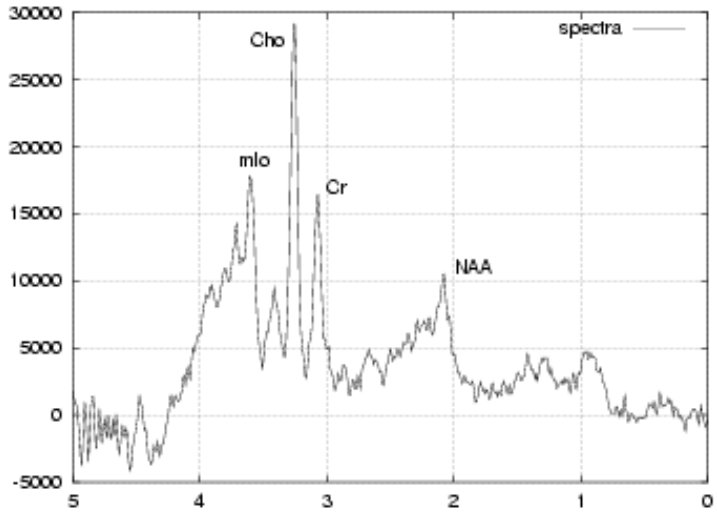

Figure 1: Typical spectra.

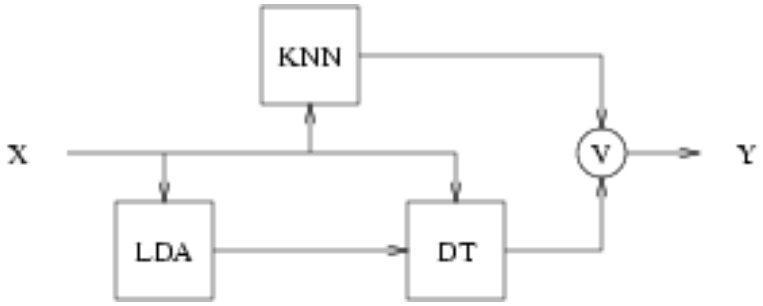

Figure 2: Cascading and voting architecture.

RESULTS

Figure 3 shows the hierarchy structure and the results of each stage of the classifier, using N-fold cross validation with $N = 10$ and averaging the results of five experiments. Precision is the percentage of samples classified as a class that really belong to such class (β), "classified samples" is the percentage of samples classified (P). Notice that P decreases as we descend the tree, as more samples are classified as "unknown", while β is usually better for the most populated classes.

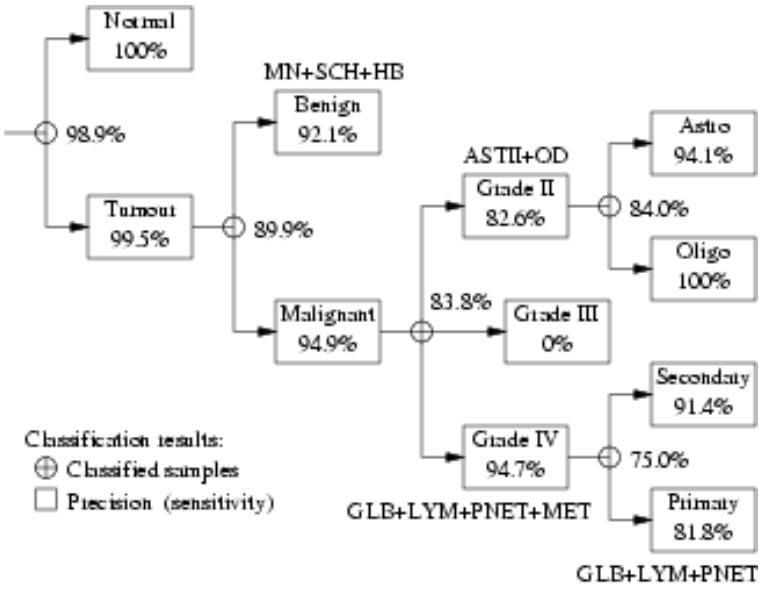

Figure 3: Classification accuracy for each category.

DISCUSSION

Although results are preliminary, and will need to be thoroughly tested with independent data, this shows a promising method for tumour classification in "real-life" clinical practice. The inherent hierarchical structure of the tumour classification problem is well described and allows all tumour types to be represented, not just the more common ones. The "unknown" class provides a simple way to increase classification performance reducing the percentage of classified samples.

References

[1] A. R. Tate et al.. Automated classification of brain tumours from 1H MRS spectra in INTERPRET, a multi-centre collaboration. In *ISMRM*, 572, 2002. `http://carbon.uab.es/INTERPRET/cdap.html` *** MORE BIB ENTRIES?