Binary Classification of Brain Tumours Using a Discrete Wavelet Transform and Energy Criteria

Carlos Arizmendi, Alfredo Vellido, Enrique Romero
Department of Computer Languages and Systems
Technical University of Catalonia,
Barcelona, Spain
e-mail: carlos.julio.arizmendi@upc.edu, {avellido, eromero}@lsi.upc.edu

Abstract— **The accurate diagnosis of human brain tumours is a sensitive medical task, for which radiology experts often must rely on indirect signal measurements. There is thus a need for developing computer-based decision support tools to assist doctors in their diagnostic task. The experiments in this brief paper address such problem in the form of binary classification, for which the pre-processing of the Magnetic Resonance Spectroscopy (MRS) signal is a most relevant data analysis stage. A combination of the Discrete Wavelet Transform (DWT) for signal decomposition and an energy criterion for signal reconstruction is used to pre-process the MRS data prior to the feature selection and classification with Bayesian Neural Networks.**

Keywords: Wavelets, Bayesian Neural Networks, MRS.

I. INTRODUCTION

Diagnosis in neuro-oncology often resorts to measurements acquired through non-invasive exploration techniques of imaging and spectroscopy. Radiology experts, therefore, must rely on their clinical experience, but also on indirect information that requires the use of computational methods for signal and image processing. These often take the form of computer-based medical decision support systems (MDSS).

The current gold standard for diagnostic classification of brain tumours is class labeling according to the World Health Organization (WHO) system, based on the histopathological analysis of biopsy samples. Biopsies require an invasive procedure that carries a risk of mortality of 0.2-0.8%, and an estimate of morbidity in the range 2.4-3.5% [1, 2]. Additionally, only about a 91% of cases are truly identifiable through this test, which means that up to 9% of patients remain undiagnosed [3]. This reinforces the importance of developing at least semi-automated MDSS tools based on non-invasively acquired information.

This study addresses the problem of human brain tumours diagnosis on the basis of biological signal data obtained by MRS. *In vivo* MRS enables the quantification of metabolite concentrations non-invasively, thereby avoiding serious risks of brain damage. We analyze a set of MRS data from the multi-centre, international INTERPRET database [4]. Previous studies have shown that the pre-processing of MRS signal is a very relevant data analysis step, strongly influencing the classification results. Here, a DWT for the decomposition of the spectra in terms of approximation and detail coefficients is combined with an energy criterion for signal reconstruction. This is followed by dimensionality reduction (DR) through feature selection using Moving Window and Variance Analysis (MWVA) or feature extraction using Principal Component Analysis (PCA), prior to classification using Bayesian Artificial Neural Networks (ANN).

II. MATERIALS

This paper investigates a multi-centre, international database of single-voxel, proton MRS (SV-^{1}H-MRS) corresponding to several brain tumour pathologies (Table 1: nine tumour pathologies, plus abscesses and normal brain tissue). It was created under the framework of the European project INTERPRET [4]: Data were collected by CDP (Centre Diagnòstic Pedralbes, Barcelona, Spain), IDI (Institut de Diagnòstic per la Imatge, Barcelona, Spain), SGHMS (St. George's Hospital Medical School, London, UK) and UMCN (University Nijmegen Medical Center, Nijmegen, Netherlands). For the experiments reported in this paper, a set of SV-^{1}H-MRS measured at short echo time (SET: 273 patients) were analyzed, comprising a total of 512 frequency intensity values (measured in parts per million (ppm), an adimensional unit of relative frequency position in the data vector.)

TABLE I: TUMOUR TYPES IN THE INTERPRET DATABASE

Tumour Class	Number of Cases
a2: Astrocytomas, grade II	22
a3: Astrocytomas, grade III	7
ab: Brain abscesses	8
gl: Glioblastomas	86
ly: Lymphomas	10
me: Metastases	38
mm: Meningiomas grade I	58
no: Normal cerebral tissue	22
oa: Oligoastrocytomas grade II	6
od: Oligodendrogliomas grade II	7
pn: Primitive neuroectodermal tumours and medulloblastomas	9

978-1-4244-9485-9/11/$26.00 ©2011 IEEE

III. METHODS

The Continuous Wavelet Transform (CWT) of a signal *x(t)* and the mother wavelet is defined as:

$$W(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t)\,\psi\left(\frac{t-\tau}{s}\right) dt \qquad (1)$$

The transformed signal $W(\tau, s)$ is a function of the translation parameter τ and the scale s. The signal energy is normalized at every scale by dividing the wavelet coefficients by $1/\sqrt{|s|}$. The original signal can be reconstructed with the inverse CWT, defined by:

$$x(t) = \frac{1}{C_\psi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(\tau, s)\, \frac{1}{s^2}\, \psi\left(\frac{t-\tau}{s}\right) d\tau ds. \qquad (2)$$

$$C_\psi = \int_0^{\infty} \frac{|\psi(f)|^2}{f} df < \infty.$$

The spectral data analyzed in this paper are discrete in nature and they cannot be processed in a practical way using CWT. Instead, DWT [5] is implemented via an octave filter bank, as a cascade of low-pass and high-pass filters, followed by sub-sampling. Every pair of filters represents a decomposition level. The reconstruction of the original signal is possible using the synthesis filter bank where the signals are upsampled and passed through the filters of reconstruction. The reconstruction procedure, except for rounding errors, leads to the restoration of the original signal if no coefficient is altered.

The DWT processing requires the selection of an adequate mother wavelet function. In this study, Daubichie and Symlet mother wavelets were investigated. An energy criterion was subsequently employed to decide on the number of wavelet coefficients that should be retained for further analysis.

Dimensionality reduction was implemented according to two strategies: feature selection using MWVA and feature extraction using PCA (of common use in radiology data analysis). MWVA is a feature selection filter method first proposed in [6]. It is based on a combination of the *moving window* technique and the analysis of between/within group variance. It should be able to identify those spectral frequencies, or intervals of frequencies, with greater ability to discriminate between tumour types.

Feed-forward Bayesian ANNs with one hidden layer were used for classification. The networks were trained with Bayesian regularization to avoid data overfitting [7] and back-propagation, updating the weights and bias according to the Levenberg-Marquardt algorithm [8]. One run of a 5-fold cross-validation was performed for each network.

This work was supported by the Spanish MICINN under grants TIN2006-08114 and TIN2009-13895-C02-01.

IV. RESULTS

A. Mother Wavelet Selection

The first task in DWT implementation is the selection of the most adequate mother wavelet. Symlet and Daubichie mother wavelets over a wide range of orders were investigated.

The wavelet decomposition yields some redundant information, and the most significant information is to be found in the lower sub-band [9], the Energy Packing Efficiency (EPE) criterion was used to eliminate this redundancy without significant distortion in the reconstructed signal. The EPE is a ratio of the total preserved energy of a certain subband, after thresholding, with respect to the total energy [10]. Histograms of the distribution of the magnitude of the calculated coefficients were first computed and the magnitude of the coefficients was then ranked in ascending order. The histogram and the accumulated energy percentage of all the spectra was calculated, observing, for all mother wavelets, that approximately 90% of the energy concentrated in a small number of coefficients.

Different quality indexes were used to help determining the optimal wavelet for our study. The first is the *standardized energy*, obtained by retaining those coefficients yielding a total energy percentage closest to 90%, divided by the number of coefficients. Fig. 1 shows the mean of the *standardized energy* for Daubichie and Symlet mother wavelets over a range of orders. The second index is the average *mean square error* (MSE) resulting from the difference between the reconstructed and the original signals. The spectra were reconstructed by retaining the coefficients of decomposition that make up 99% of total accumulated energy. Fig. 2 shows the MSE and the number of coefficients for Daubichie and Symlet mother wavelets.

In Fig. 1, the highest standarized energy value (10.1) is obtained with a Daubichie of order 2, using 12 decomposition coefficients. The best result for Symlet is a lower 3.9, obtained for wavelets of order 1, using 23 decomposition coefficients. The decomposition based on Daubichie wavelets is therefore preferred.

Note also that the differential in MSE between the Daubichie of order 1 and the Daubichie of order 45 is a meager 1.6, while the difference in the number of coefficients is of almost 50: too high an increase of system dimensionality for such small error improvement. It could therefore be argued that the selection of a Daubichie mother wavelet of order 2 is a good compromise for retaining a high

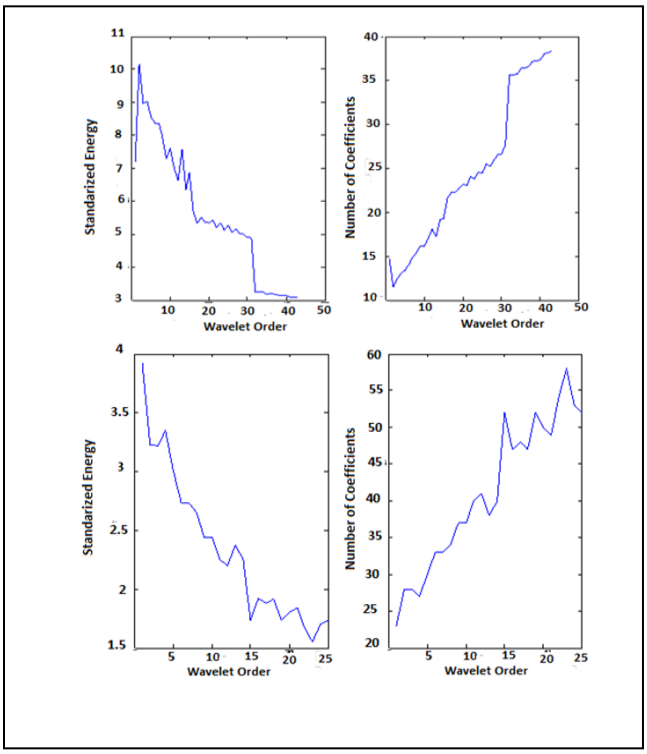

Figure 1. The top row displays correspond to the standarized energy and number of coefficients when Daubichie is computed, while the bottom row correspond to the standarized energy and number of coefficients when Symlets are computed.

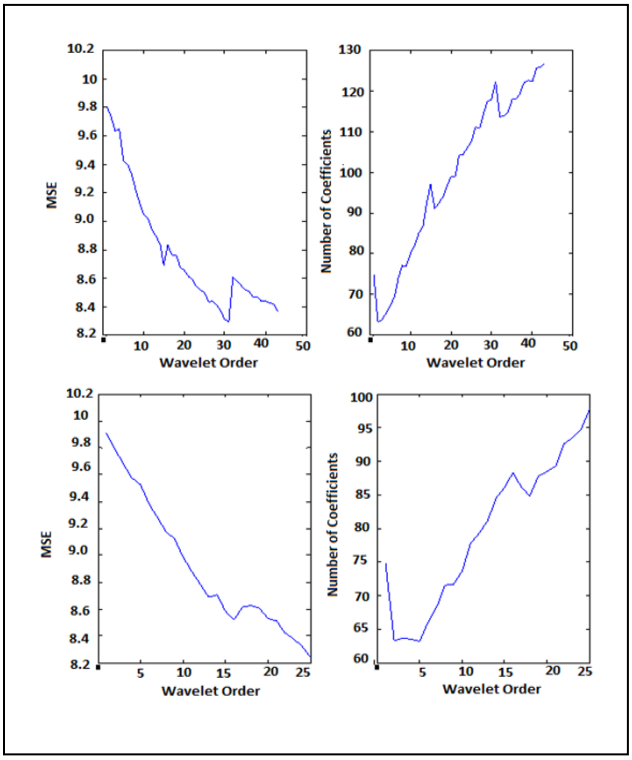

Figure 2. The top row displays correspond to the MSE and number of coefficients when Daubichie is computed, while the second bottom correspond to the MSE and number of coefficients when Symlets are computed.

standarized energy using the minimum number of coefficients for the reconstruction of the spectra, while not suffering an excessive increase of MSE.

B. Dimensionality Reduction and Classification

Once the spectra were decomposed using Daubichie wavelets of order 2, feature selection and extraction were implemented using, in turn, MWVA and PCA, starting from the obtained decomposition coefficients. For PCA, principal components were added one at a time until the differential accumulative variance between components was less than 1%, this way obtaining a 94.833% of average of variance explained.

An average of 10.2 and 10.7 features were respectively obtained for MWVA and PCA. This is a drastic reduction of dimensionality that should ease the classification task, as well as improve the interpretability of the results.

Starting from the selected and extracted features, Bayesian ANNs with one hidden layer consisting of between 5 and 40 hidden units were used in the classification experiments. The networks were trained as described in the methods section. To address the issue of class imbalance (the number of cases available from each tumour type is always small, but widely varying, as reported in Table 1), the original datasets were re-sampled, by over-sampling the minority class and under-sampling the majority class [11].

Table II summarizes, the best results of the area under the ROC curve (AUC) and the accuracy for all experiments. G1 (low grade gliomas) is the union of a2, oa and od. G2 (high-grade malignant tumours) is the union of gl and me. With only a few exceptions, the MWVA feature selection strategy yielded better accuracy results than PCA-based feature extraction. Some experiments (those involving normal tissue, as well as mm *vs.* ab) achieved perfect accuracy. Accuracy fell below 90% in only four problems, all of them well-know for their difficulty: the discrimination between astrocytomas of similar grade (a2 *vs.* a3, see, for instance [12]); the discrimination between high-grade malignant tumours (gl *vs.* me, see [13], and two problems involving the differentiation of lymphomas from high-grade tumours(gl *vs.* ly and me *vs.* ly), which have received little attention in the neuro-oncology literature.

V. CONCLUSIONS

Data pre-processing is known to be of great importance in problems of tumour type classification based on MRS signal. A combination of the Discrete Wavelet Transform (DWT) for signal decomposition and an energy criterion for signal reconstruction was used in this brief paper as a previous step to data dimensionality reduction and

classification using Bayesian ANNs. The diagnostic tumour type differentiation yielded very encouraging results using this data analysis methodology. The accurate discrimination of pathologies that have seldom been analyzed in a similar setting should be of special interest to radiology experts.

TABLE II. MEAN ± STANDARD DEVIATION OF AUC AND ACCURACY VALUES FOR ALL THE EXPERIMENTS

	MWVA	***PCA***	***MWVA***	***PCA***
Experiments	***AUC***		***ACCURACY***	
G1 *vs* G2	0.95±0.05	0.95±0.05	91.73±4.86	89.544±5.41
G1 *vs* mm	0.99±0.02	0.96±0.04	94.75±5.55	89.75±7.15
a2 *vs* a3	0.92±0.11	0.96±0.04	88.00±10.95	64.00±21.90
a2 *vs* G2	0.99±0.02	0.96±0.05	96.73±3.36	91.80±0.18
a2 *vs* ly	1.00±0.00	1.00±0.00	96.00±8.94	88.66±10.43
a2 *vs* oa	1.00±0.00	1.00±0.00	96.00±8.94	66.00±13.41
a3 *vs* pn	1.00±0.00	0.80±0.00	93.33±14.90	73.33±27.88
G2 *vs* mm	0.97±0.02	0.97±0.02	92.71±4.39	94.10±2.66
gl *vs* a3	0.91±0.12	0.98±0.01	93.50±4.71	91.08±5.50
gl *vs* ab	0.95±0.08	0.88±0.06	94.91±2.84	84.91±10.37
gl *vs* ly	0.94±0.06	0.94±0.05	87.50±4.41	90.00±12.18
gl *vs* me	0.75±0.11	0.66±0.12	76.04±8.72	55.90±8.56
gl *vs* no	1.00±0.00	1.00±0.00	100.00±0.00	96.66±3.04
gl *vs* pn	0.98±0.03	0.94±0.06	91.25±7.12	87.50±7.65
me *vs* ly	0.96±0.03	0.93±0.06	87.50±8.83	75.00±12.5
me *vs* mm	0.99±0.01	0.94±0.04	93.75±0.00	87.50±6.25
me *vs* no	1.00±0.00	1.00±0.00	100.00±0.00	96.00±8.94
me *vs* pn	1.00±0.00	1.00±0.00	95.00±6.84	87.50±8.83
mm *vs* ab	1.00±0.00	1.00±0.00	100.00±0.00	94.54±8.13
od *vs* a2	1.00±0.00	0.95±0.06	92.00±10.95	60.00±28.28

ACKNOWLEDGMENT

Authors acknowledge former INTERPRET European project partners. Data providers: Dr. C. Majós (IDI), Dr, À. Moreno-Torres (CDP), Dr. F.A. Howe and Prof. J. Griffiths (SGUL), Prof. A. Heerschap (RU), Dr. l. Stefanczyk and Dr. J. Fortuniak and Dr J. Calvar (FLENI); data curators: Dr. M. Julià-Sapé, Dr. A.P. Candiota, Ms. T. Delgado, Ms. J. Martín, Dr. I. Olier and Mr. A. Pérez (all from GABRMN-UAB). Prof. C. Arús, GABRMN coordinator.

References

[1] J. Favre, J.M. Taha, KJ Burchiel. An analysis of the restrospective risk of hematoma formation in 361 consecutive morphological and functional stereotactic procedure. *Journal of Neurosurgery*. 2002; 50(1): 48-56;discussion 56-57.

[2] M. Field, T.F. Witham, J.C. Flickinger. D Kondziolka, LD Lunsford. Comprehensive assessment of hemorrahage risks and outcomes after stereotactic brain biopsy. *Journal of Neurosurgery*. 2001;94(4);545-551.

[3] W.A. Hall. The safety and efficacy of stereotactic biopsy for intracranial lesions. *Cancer*. 1998; 82(9): 1749-1755.

[4] M. Julià-Sapé, D. Acosta, M. Mier, C. Arús, D. Watson, & the INTERPRET Consortium. A multi-centre, web-accessible and quality control checked database of in vivo MR spectra of brain tumour patients. Magnetic Resonance Materials in Physics. *MAGMA*, 2006, 19, 22–33

[5] S. Mallat. A *Wavelet Tour of Signal Processing,* Usa: Academic Press, 1999.

[6] C. Arizmendi, A. Vellido, E. Romero . Frequency selection for the diagnostic characterization of human brain tumours*, Frontiers in Artificial Intelligence and Applications*, 2009, 202, 391-398,.

[7] D. MacKay. The evidence framework applied to classification networks*. Neural Computation*, 1992, 4 (5); 720–736.

[8] F. Foresee, and M.T. Hagan. Gauss-Newton approximation to Bayesian regularization, Proceedings of the 1997 International Joint Conference on Neural Networks, 1997, 1930-1935.

[9] T.I. Mohammadpour, M.R. Karami. ECG Compression with Thresholding of 2-D Wavelet transform coefficients and run length coding, *European Journal of Scientific Research*, 2009, 27, pp. 248-257

[10] B.A. Rajoub. An Efficient Coding Algorithm for the Compression of ECG Signals Using the Wavelet Transform, *IEEE Transaction. in Biomedical Engineering,* 2002 49(4);. 355-362.

[11] N. Japkowicz. The Class Imbalance Problem: Significance and Strategies. In Proceedings of the 2000 International Conference on Artificial Intelligence (IC-AI'2000):Special Track on Inductive Learning Las Vegas, Nevada

[12] C. Ladroue. *Pattern Recognition Techniques for the Study of Magnetic Resonance Spectra of Brain Tumours*, PhD thesis, St. George's Hospital Medical School, 2003.

[13] E. Romero, A. Vellido, M. Julià-Sapé, and C. Arús. Discriminating glioblastomas from metastases in a SV 1H-MRS brain tumour database. In Proceedings of the 26th Annual Meeting of the European Society for Magnetic Resonance in Medicine and Biology (ESMRMB), 2009.