ISSN:0975-9646
Meghana Nagori et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (5) , 2015, 4829-4832

Comparative Analysis of Classifiers for Brain Tumour Detection Using Metabolites from Magnetic Resonance Spectroscopy Graph

Meghana Nagori[1], Dr. Madhuri Joshi[2]

[1]*Research Scholar,Department of Electronics and Telecommunications,,Government Engineering College Railway Station Road,Aurangabad,Maharashtra,India*

[2]*Professor,Jawaharlal Nehru Engineering College Cidco,Aurangabad,Maharashtra,India*

***Abstract*— Last decade has seen lots of research in automatic brain tumour classification by using metabolite lipids like N-Acetylasparate, Creatine and Choline values from Magnetic Resonance Spectroscopy (MRS) graphs. Accurate identification of the type of the brain abnormality is highly essential since the treatment planning is different for all the brain abnormalities. Any false detection may lead to a wrong treatment which ultimately leads to fatal results. Early detection and diagnosis of brain tumour can reduce fatality to an extent. This research aims at supporting clinicians in preliminary decision making. It does the same by improving the accuracy of the classifiers by significantly reducing false positives .The classifiers employed are Naïve Bayes Classifier, Random Tree and Instance based classifier. The maximum classification accuracy of around 95%-100% is achieved.**

***Keywords*- MRS, Metabolites, Classifiers, Confusion Matrix, Naïve Bayes**

I. INTRODUCTION

MRI spectroscopy (MRS) offers the capability of using MRI to noninvasively study tissue biochemistry. In MRS, either 1H atoms in other molecules or other atoms such as 31P, 23Na, K, 19F, or Li are flipped. Within a given brain region called a voxel, information on these molecules is usually presented as a spectrograph with precession frequency on the x-axis revealing the identity of a compound and intensity on the y-axis which helps quantify the amount of a Substance.

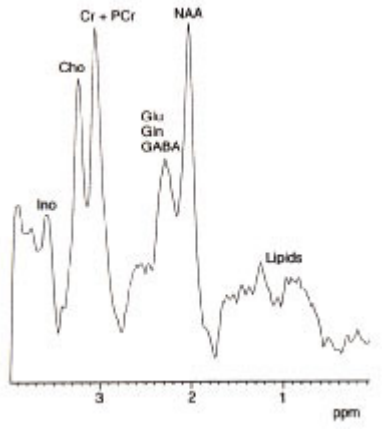

Fig.1 MRS Spectroscope

The quantity of a substance is related to the area under its spectrographic peak; the larger the area, the more of a substance that is present.

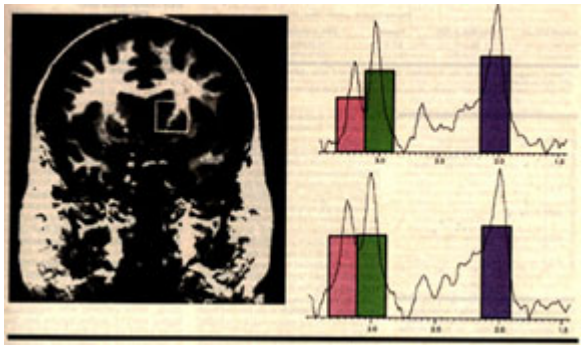

Fig. 2 fMRI image and its equivalent spectroscope

MRS can be used to identify regional biochemical abnormalities. MRS can also be used to measure changes in metabolic activity in neurology and neurosurgery, however, MRS is starting to be used in the characterization of tumour. Magnetic Resonance (MR) brain image classification is a mandatory but challenging task in the medical field. Accurate identification of the nature of the disease is highly essential for the successful treatment planning. One of the significant applications of image classification is the medical field in which the abnormal brain tumor images are categorized prior to treatment planning. Accurate identification of the type of the brain abnormality is highly essential since the treatment planning is different for all the brain abnormalities.

II. RELATED WORK

Brain tumour classification has been performed using long echo proton MRS signals. [1]The major limitation is the limited number of available spectra for the tumour types which results in inferior classification accuracy. Brain tumour classification has also been implemented using wavelets. But the major drawback is the low convergence rate. Expectation-maximization techniques are also used for brain tumour classification. But the major limitation is the requirement of a spatial probabilistic atlas that contains expert prior knowledge about the brain structures. J. L. Griffin and R. A. Kauppinen[2] discuss in their paper a technique called

HMRS which is used to identify brain tumour cells. A comprehensive list of metabolites that are used in detection of brain tumour is discussed by the authors. Other than HMRS,PMRS and CMRS metabolite content can also be used in distinguishing brain tumour cells.[3] Martorell, Olier, Julià-Sapé, and Arús[4] have developed a fully automated classification system for MRS data. PCA is the method that is used for feature extraction and FDA is the method selected for classification.. The project INTERPRET has allowed pattern recognition software to aid clinicians with the differentiation of brain tumours using MRS. [5]

III. Proposed Architecture

The architecture of the proposed system includes the following modules:-

Image Processing:-

Different methods are designed for extracting metabolite values from MRS graph images and storing it in the form of numbers.

Feature Extraction and Pattern Recognition:-

A feature vector set from the extracted metabolite values is created to develop a pattern for each category of tumour class.

Machine Learning:-

Various classifiers are employed for categorizing the patterns into given tumour classes.

Statistical Analysis:-

A comprehensive analysis employing different performance metrics is carried out to test the accuracy of different models.

Results:-

It displays the tumour class for each patient.

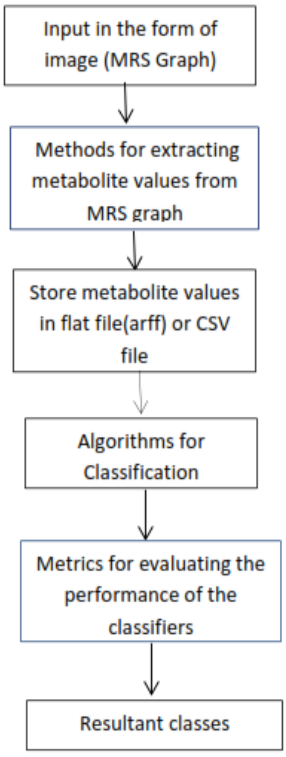

Fig. 3 Flow of the proposed system

As shown in fig.3 the MRS images were collected from four prominent data sources .Relevant methods of extracting metabolite values from MRS graph were developed, as storage of graphical images incurs lots of space. The values were stored in flat file format. After that either CSV file itself or an arff file acted as input to the various classification algorithms. The results were evaluated by different performance metrics like sensitivity and specificity. The result is the number of patients identified and classified into the three classes –benign, malignant and infection.

IV. Dataset Details

A. MATERIALS AND METHODS

For the dataset under consideration, in each patient, axial T1-weighted, T2-weighted FSE, and fast fluid attenuated inversion recovery (FLAIR) images were obtained. After intravenous administration of gadolinium contrast-enhanced T1 weighted SE sequences were obtained in axial, coronal, and sagittal planes. Multi echo multi planar sequences through the brain were carried out both before and after administration of IV contrast. Multi-voxel spectroscopic data were obtained in all patients. As data was collected from multiple sources the scanning parameters for some images differed and for T2 sagittal and coronal's and T1, T2, GRE and FLAIR were obtained. Diffusion DW images and single-voxel spectroscopy (35/144 TE) were obtained for some instances. A study population comprising of one hundred and twenty two patients collected from four different hospitals are evaluated using metabolites present in MRS graph.

B. ATTRIBUTES USED FOR EXPERIMENTATION

Table 1. below depicts the various attributes used in experimentation.

Table. 1 Attributes and their characteristics

Attributes	Attributes Description	Attributes Usage
Cho	Choline	Used for identification of wide range of brain tumours
NAA	N-Acetyl Asparate	Decreases in brain tumours
Cr2 and Cr	Creatine	Decreases in brain tumours
NAA/Cho	N- Acetyl Asparate Choline	Opposite of Cho/NAA ratios
Cho/NAA	Choline/N- Acetyl Asparate	Increases in case of tumour
NAA/Cr	N- Acetyl Asparate/Creatine	Decreases in brain tumours

V. PREPROCESSING TECHNIQUES

The methods employed for extracting the values of the metabolites from the graph were: [6]

A. Value extraction from most visible peak point

The MRS graph consists of peaks namely Cr2, Cho,Cr and NAA. This method extracts one of the most clear and easily readable peak points. Same peak point is projected on Y-

axis and the scale value is taken which is nearer to it. The scale is approximately calibrated. Once the scale is defined other peak points which are not clearly readable are projected on Y axis and the ordinate value is taken. The peak point values calculated as the multiplication of scale for 1 unit and ordinate value. In this approach, the peak points are calculated approximately.

B. Graph Scanning

From the MRS graph the noise is first removed and then after converting it into binary image edge detection algorithms are applied to find peak values. This leads to identification of regions of interest from input graph and then by employing OCR the peak values are extracted and stored in the form of numbers in the database.

VI. Classification Techniques

A. Random Tree

In this method, instead of using all the attributes (in our case all the metabolite ratios) a subset is used first and the tree is constructed..In this way by using random features at each node a set of random possible trees gets created which leads to equal chance of being sampled. Thus the combination of large sets of random trees leads to generate models efficiently and accurately [7]

B. Naïve Bayes

Bayesian Probability theory is used to judge the truth of the hypothesis of the given data.[8] The truth of the Hypothesis is carried out by checking some combination of values of attributes which classifies the instance to some class .This is carried out using Bayes rule:-

$$P(H/D) = P(D/H)\, P(H)/P(D) \quad (1)$$

Here D is instance that we which to check
H is hypothesis
And term P (D/H) is likelihood function.
This is a statistical method of learning and it reflects what we have learned about the validity of the hypothesis by considering the data.

C. IB1

These learning algorithms employ a variant of the k-nearest neighbour pattern classifier and can learn using a polynomial number of instances .They can also be used to generate classification predictions using only specific instances as in our case. These algorithms use similarity functions to yield graded matches between instances. They can also focus on theory-based reasoning in real world applications. It has been proved that such algorithms perform well in real-world domains. [9]

VII. Performance Evaluation and Results

The classifiers are evaluated by confusion matrix. The model performance can be assessed by using the metrics true positive rate(sensitivity),false positive rate(specificity), precision, f-measure, ,accuracy and kappa statistics. The words sensitivity and specificity have their origins in screening tests for diseases. Sensitivity is defined as the probability that the test says a person has the disease when in fact they do have the disease.

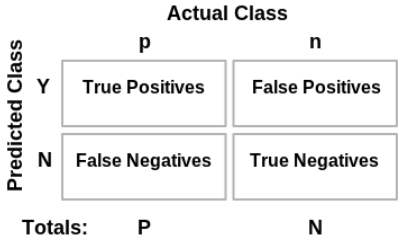

Fig. 4 Confusion Matrix

$$\text{Sensitivity} = TP/TP+FN \quad (2)$$

Specificity is defined as the probability that the test says a person does not have the disease when in fact they are disease free.

$$\text{Sensitivity} = A/A+B \quad (3)$$
$$\text{Accuracy (AC)} = A+D/A+B+C+D \quad (4)$$
$$\text{Error rate} = 1-AC \quad (5)$$
$$\text{Specificity} = D/C+D \quad (6)$$
$$\text{Precision} = A/A+C \quad (7)$$
$$\text{F-Measure} = P+R/2PR \quad (8)$$
$$\text{Where P} = \text{Precision} = \text{Recall (Sensitivity)} \quad (9)$$

The values computed are between 0 and 1 and the larger f-measure values indicate good classification quality.
Since the class distribution is unbalanced, it is important to use a performance measure that takes class imbalance into account. The authors use the measure of Kappa Statistic due to Cohen [11] .The value of kappa statics is given by:--

$$k = D\text{observed} - D\text{random}/D\text{perfect} - D\text{random} \quad (16)$$

where D is the diagonal entries from the confusion matrix.
A value of $k>0.8$ demonstrates a good reliable classifier whereas a value of $.67<k<.8$ may indicate tentative conclusions by the classifier. The common way to analyse clinical spectra is to look at metabolite ratios, namely NAA/Cr, NAA/Cho, and Cho/Cr. Normal and abnormal values are shown in the table below.

Table 2 Metabolite ratios

Metabolite ratio	Normal	Abnormal
NAA/Cho	1.6	<1.2
NAA/Cr	2.0	<1.6
Cho/Cr	1.2	>1.5

It was found out that nearly all brain tumours have decreased N-acetyl aspartate(NAA) signals ,and often have increased levels of choline(Cho) leading to increased Cho/NAA ratios.[12] Increased Cho signal (evaluated as Cho levels Cho/Cr or CHO/NAA)are suggestive of recurrences while significantly reduced Cho and Cr levels are suggestive of radiation necrosis. When Cho/NAA ratio is greater than 1.9 there is strong indication of presence of

tumour. Tables 3-6 display the results of experimentation on different classifiers.

Table 3 Confusion matrix for Naïve Bayes

A	B	C	Classified as:-
44	2	2	A=malignant
1	44	0	B=benign
0	0	29	C=infectious-diseases

Table 4 Confusion matrix for IB1

A	B	C	Classified as:-
47	1	0	A=malignant
1	44	0	B=benign
0	0	29	C=infectious-diseases

Table 5 Accuracy of algorithms

Algorithm	Accuracy
Random Tree	100%
Naïve Bayes	95.9016%
IB1	98.3607%

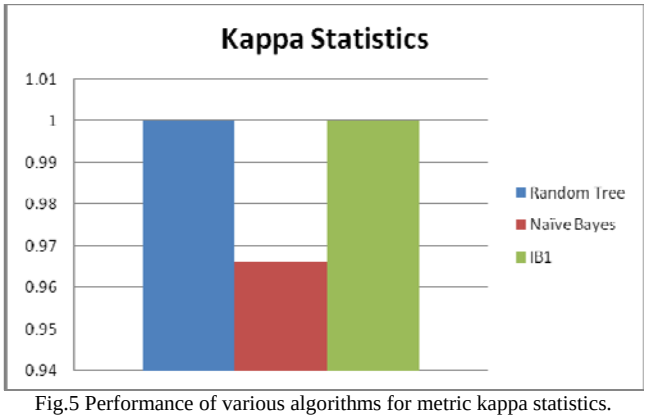

Fig.5 Performance of various algorithms for metric kappa statistics.

VIII. CONCLUSIONS

MRS can be valuable in the initial evaluation of brain tumours because of the different patterns of metabolites seen in tumours and because the abnormalities become more severe in advanced disease. In most of the Hospitals, the MRS graphs of brain tumour patients are scanned and then the observed metabolite values are manually entered into database. This process is time-consuming. There are disadvantages while working with the images like large storage, incurs more time in opening and closing and corrections are not allowed. Hence there is the need to automate the process. The approaches for extracting metabolite values from MRS graph require less storage than a graph stored as an image format. This study aims at diagnosing and identifying brain tumour disease pattern by automating extraction procedure of metabolite values and then applying machine learning based methods of random tree, Naive Bayesian, and IB1.On the sample dataset the authors obtained 100% accuracy in tumour disease pattern identification by random tree and around 95% on average by other methods. This result may suggest that these methods can be used confidently for other diagnosis problems too. Till to date there is no criteria for separating all tumour types from MRS. Finally MRS can be regarded as having a great role in predicting/managing the disease before it becomes life threatening.

REFERENCES

[1] Alena Horská, Peter B. Barker," Imaging of Brain Tumors: MR Spectroscopy and Metabolic Imaging", *Neuroimaging Clin N Am*,.vol.20(3),pp. 293–310, 2010

[2] J.L.Griffin,R.A.Kauppinen,"A metabolomics perspective of human brain tumours",*FEBS*,vol. 274(5),pp.1132-1139,2007

[3] Alessandro Alimenti,Jacqueline Delavelle,François Lazeyras, Hasan Yilmaz ,Pierre-Yves Dietrich, "H Magnetic Resonance Spectroscopy in the Progression of Gliomas",.*European neurology*

[4] Martorell,Olier,Sape,Arus,"Automatic relevance source determination in human brain tumors using Bayesian NMF",in *Proc CIDM*,2014,p..99-104

[5] Alexander Pérez-Ruiz, Margarida Julià-Sapé, Guillem Mercadal, Iván Olier, CarlesMajós, CarlesArús," The INTERPRET Decision-Support System version 3.0 for evaluation of Magnetic Resonance Spectroscopy data from human brain tumours and other abnormal brain masses", *BMC Bioinformatics*, vol.11,pp 1-12,2010

[5] D.JudeHemanth, C.KeziSelvaVijila,J.Anitha," Performance Improved PSO based Modified Counter Propagation Neural Network for Abnormal MR Brain Image Classification",.*Int. J. Advance. Soft Comput. Appl*,vol .2(1),2010

[6] Meghana Nagori,Madhuri Joshi, "Methods and Algorithms for Extracting Values from MRS Graph for Brain TumourDetection",in *PROC EECS*,2013,p. 331-336

[7] Breiman, L., "Random Forests", *Machine Learning*,vol.45(1),pp. 5-32,2001.

[8] Igor Kononenko, "Inductive and Bayesian Learning in Medical Diagnosis", *Proc* 1st *Congress* on Yougoslav *Medical* Informatics

[9] David Aha,dennis kibler,Marc k Albert,"Instance based learning algorithms",*Machine learning*,vol..6,pp. 37-66,1991

[10] J. Cohen, "A coefficient of agreement for nominal scales", *Educational and Psychological Measurement*,vol.20(1),pp.37–46,1960.

[11] Mohamed Fouad Sherif , Faten Mohamed Salem , Magdy A. Almahallawy ,Ahmad Mohamed Abd Algawad , Qanet M. Hammad , "Role of magnetic resonance spectroscopy in differentiation between recurrence of glioma and post radiation injury", *The Egyptian Society of Radiology and Nuclear Medicine*,vol 45,pp.1233-1240,2014