ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, January 2013

Detection of Brain Tumor by Mining fMRI Images

Meghana Nagori[1], Shivaji Mutkule[2], Praful Sonarkar[3]

Asst. Professor, CSE, Government Engineering College, Aurangabad, India [1]

ME Student, CSE, Government Engineering College, Aurangabad, India [2, 3]

Abstract: Brain tumor patients are increasing day-to-day. This paper proposes a novel approach to extract metabolite values from graph. Metabolites like NAA, Creatine, Choline and Cr2 are used to detect the brain tumor. Cho/NAA ratio plays most important role in deciding the tumor type so weights are assigned to each metabolite while clustering. Clustering algorithms could able to achieve accuracy up to 86%. Proposed system is based on decision tree algorithms which are proven to be better against clustering algorithms. Proposed system stores the metabolite values in dataset instead of storing fMRI images so reduces the image processing tasks and memory requirements.

Keywords: fMRI, Clustering algorithms, Classification algorithms, Z-score ranking, K-means, FT, C4.5 algorithm.

I. INTRODUCTION

Brain tumor is abnormal and uncontrolled growth of brain cells. Functional Magnetic Resonance Imaging (fMRI) provides detail information about brain tumor anatomy, cellular structure, making it an important tool for diagnosis [1].

372 people were diagnosed with Brain and Central Nervous System tumors in TATA Memorial Hospital in Mumbai, India. Out of 372, 250 were males (67%) and 122 were females (33%) [2].

National Cancer Institute estimates 22,070 new cases and 12,920 deaths in the US for 2009 [3]. According to Central Brain Tumor Registry of US (CBTRUS), 64,530 new cases of primary brain and central nervous system tumors diagnosed at the end of 2011 [4].

Estimated new cases are 23,130 and deaths are 14,080 from brain and other nervous system cancers in the United States in 2013 [5].

Sample MRI image is shown in Fig. 1 with four types of metabolites namely N-acetylaspartate (NAA), Creatine (Cr), Choline (Cho) and Methylene protons of creatine (Cr2). In proposed system metabolite values are extracted and computed Cho/NAA ratio which plays important role in brain tumor detection.

There is huge data in hospitals for brain tumor patients. This data is present in fMRI image format. This dataset can be mined to get knowledge. In proposed system metabolite values are extracted from fMRI images. Extracted values are stored in dataset. Memory requirement for brain tumor dataset is much less than fMRI images. Generated dataset is subjected to various clustering and classification techniques.

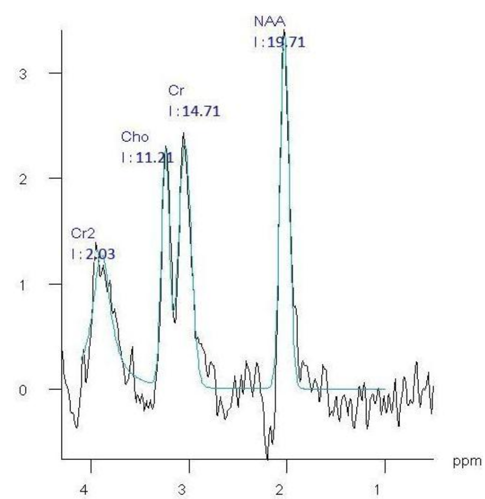


Fig. 1 Sample fMRI Graph Image

II. LITERATURE SURVEY

P. Rajendran and M. Madheswaran in [6] proposed association rule mining technique to classify the CT scan brain images. For this study three categories have been taken namely normal, benign and malign. Low level feature extracted from images and high level knowledge from specialists is combined into system.

hasan aydin, nilay aydin oktay, serdar spaholu, elif altin and baki hekmolu in [7] evaluated proton MR spectroscopy

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, January 2013

for brain tumor categorization. Brain tumors are classified into low-high grade glial neoplasms, menengiomas and metastasis. Brain tumors are categorized on the basis of Cho/NAA, Cho/Cr and Cho/MI metabolite ratios.

T. Logeswari and M. Karnan in [8] proposed segmentation based brain tumor detection. Proposed segmentation method has two phases. In the first phase, film artifact and noise are removed. In second phase, Hierarchical Self Organizing Map (HSOM) is applied.

G Vijay Kumar and Dr GV Raju in [9] proposed early prediction of brain cancer based on texture features and neuro classification logic. Nine distinct features along with minimum distance are used for brain tumor detection in given MRI image. Extracted region is recognized using neuro fuzzy approach.

Andac Hamamci, Nadir Kucuk, Kutlay Karaman, Kayihan Engin, and Gozde Unal in [10] presents fast and robust tool for segmentation of solid brain tumors. Tool assists clinicians and researchers in radio surgery planning with minimal user interaction.

Dina Aboul Dahab, Samy S. A. Ghoniemy and Gamal M. Selim in [3] applied modified segmentation techniques on MRI scan images to detect brain tumor. Modified Probabilistic Neural Network based on Learning Vector Quantization with image and data analysis to classify brain tumor using MRI scans.

Sudipta Roy and Samir K. Bandyopadhyay in [11] proposes fully automatic algorithm for brain tumor detection using symmetry analysis. Disease progression is indicated by quantitative analysis.

III. PROPOSED SYSTEM

Proposed system takes fMRI images as an input. Extract values from fMRI graph and store it in dataset. Dataset can be used in clustering or classification. Overall working of proposed system is shown in Fig. 2.

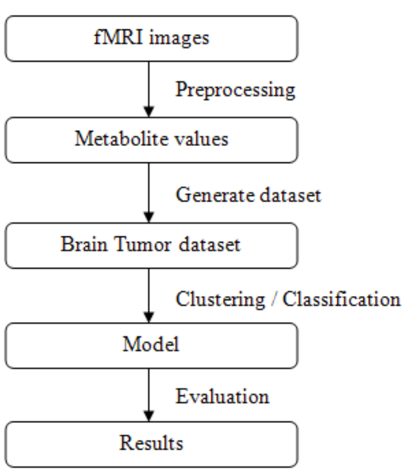


Fig. 2 Proposed System Design

A. *Preprocessing*

fMRI images contains metabolite values those can be used to detect the brain tumor type. Simple graph scanning method is used to extract the values from graph. Graph scanning algorithm is shown in Fig. 3.

1. Locate X-axis of MRI image. (X-axis is same for all MRI images i.e. in PPM with same scale)
2. Metabolite[0]=Location of 2 on X-axis as NAA metabolite always occurs at 2 PPM.
3. Metabolite[1]= Location of 3 on X-axis as Cr metabolite always occurs at 3 PPM.
4. Metabolite[2]= Location of peak next to Cr.
5. Metabolite[3]= Location of peak less than but near to 4PPM.
6. Take suitable scale value *s* from user.
7. For each *m* in Metabolite
   a) Scan along Y- axis until peak point is detected on MRI graph line.
   b) Compute distance between peak point and X axis.
   c) Temp=distance / 115.
   d) Result[*m*]=Temp × *s*.
8. End
9. Store Result array.
10. End

Fig. 3 Graph Scanning Algorithm

The graph scanning algorithm is about detecting the peak points from graph and noting the co-ordinate values. The distance between peak point and X-axis is calculated. The calculated distance is multiplied with suitable user scale to get the metabolite value nearer to the original one. Extracted values are stored in excel sheet in Fig. 4.

	A	B	C	D	E	F
1	CHO/NAA	CHO	NAA	CR2	CR	Result
2	0.51	10	19.7	4.61	11.22	Benign
3	0.67	14.59	21.71	7.59	13.16	Mild
4	1.16	3.9	3.37	3.63	2.78	Infection
5	1.42	14	9.84	5.46	6.46	Malignant
6	0.76	9.03	11.82	6.13	10.91	Mild
7	0.72	9.3	12.97	6.34	8.32	Mild

Fig. 4 Snapshot of Excel Sheet

B. *Clustering and Classification Algorithms*

Three algorithms are considered for brain tumor detection namely

1. Weighted K-means with Z-score ranking method
2. Functional trees algorithm
3. J48graft algorithm

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, January 2013

First algorithm is unsupervised learning and last two are supervised learning algorithms. These three algorithms can be briefly described as

1. Weighted K means with Z score ranking method

Accuracy of K-means algorithm depends on initial centroid selection and order of instances. This drawback is removed using Z-score ranking method to re-order the instances and select initial centroids [12]. Same algorithm does not work well in brain tumor detection because simple K-means with Z-score assigns same importance to all attributes *i.e.* metabolites. But in brain tumor detection the attribute Cho/NAA has more importance than other. So weight is assigned to each metabolite using equation 1.

$$W = \frac{1}{\sigma^2} \quad (1)$$

Where W is weight of attribute and σ is the standard deviation of attribute. Standard deviation can be calculated using equation 2.

$$\sigma = \sum_{i=1}^{n} \sqrt{\frac{(x_i-\mu)^2}{n-1}} \quad (2)$$

Where μ is mean value as

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n} \quad (3)$$

Calculated weights are used in Z-score formula to provide the importance to each attribute. As per the need, one may choose some arbitrary weight to each attribute. The statistical Z-score ranking formula with weight is given by

$$Z\ score = \frac{(x_i-\mu)\times W_i}{\sigma} \quad (4)$$

So using above Z-score formula the data is ranked with more importance to specified attribute.

Proposed algorithm is combination of weighted K-means and Z-score ranking method. Not only Z-score formula is redefined but also distance formula is redefined as

$$d = \sqrt{(x_1-x_2)^2 \times W_x + (y_1-y_2)^2 \times W_y} \quad (5)$$

The proposed algorithm is described in Fig. 5

- Input: *K*-number of clusters, D-dataset with n objects.
- Output: A set of *K* clusters.

1. Calculate Initial centroids using Z-score as
   a) Calculate weighted Z-score of each point using equation 4.
   b) Sort dataset based on Z-score values.
   c) Divide dataset into *K* subsets.
   d) Calculate mean value of each subset.
   e) Take any data point value as initial centroid which is nearer to calculated mean.
2. (Re) assign each data point to cluster based on similarity. Similarity is calculated by equation 5.
3. Update cluster mean as data point may move from one cluster to another.
4. Go to step (2) until no change.

2. Functional trees algorithm

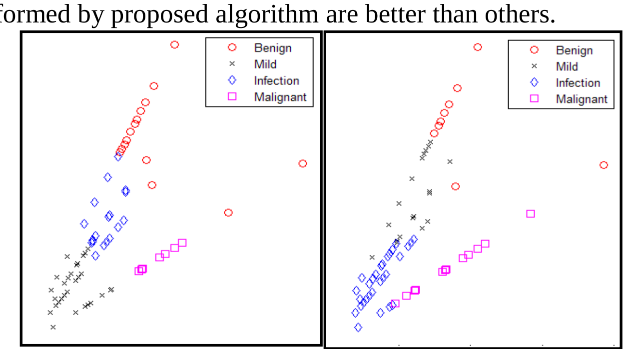
Functional trees are implemented in weka. Functional trees are classification tree that could have logistic regression function at the inner node and/or leaves. Algorithm can deal with numeric and nominal attributes with missing values [13][16].

3. J48graft algorithm

J48graft algorithm generates grafted C4.5 decision trees. Tree grafting technique increases predictive accuracy of a classifier [14-16].

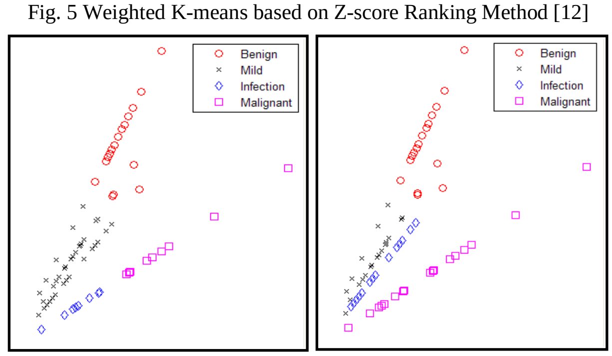
IV. EXPERIMENTAL RESULTS

All the above algorithms are applied on brain tumor dataset generated using graph scanning method. The dataset consist of 76 instances out of which 21 instances are Benign, 22 instances are Mild, 19 instances are Malignant and 14 instances are Infection. Performance of Weighted K-means with Z-score ranking algorithm is compared with simple K-means, weighted K-means and K-means with Z-score. Weighted K-means with Z-score performs better with respect to other clustering algorithms. Clustering graphs for different flavours of K-means are shown in Fig. 6. Cho is considered along X-axis and NAA is along Y-axis. Clusters formed by proposed algorithm are better than others.



(a) Simple K means (b) Weighted K means

Fig. 5 Weighted K-means based on Z-score Ranking Method [12]



(c) K means with Z score (d) Weighted K means with Z score

Fig. 6 Different flavours of K-means are plotted with Cho Vs NAA

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 4, January 2013

The accuracy of each algorithm is calculated with respect to changing data instances. Proposed weighted K-means with Z-score performs well with respect to others as shown in Fig. 7.

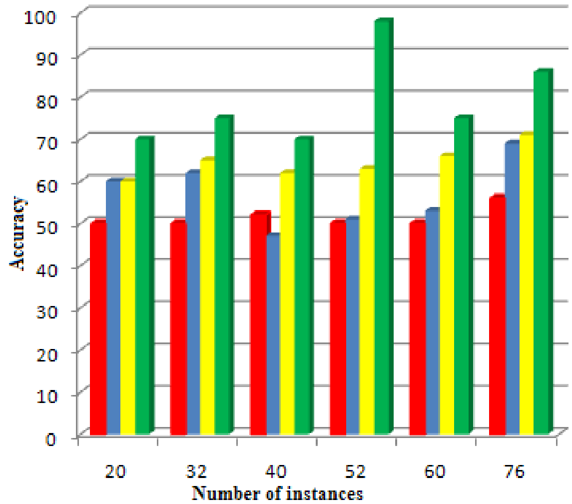


Fig. 7 Accuracy of simple K-means, weighted K-means, K-means with Z-score and weighted K-means with Z-score respectively

The accuracy weighted K-means with Z-score algorithm is compared with other supervised learning algorithms. So, we can conclude that whether unsupervised or supervised is better for detection of brain tumor. The graph is drawn in Fig. 8 for False Positive Rate, Precision, True Positive Rate, Area under Curve and Accuracy for 76 instances.

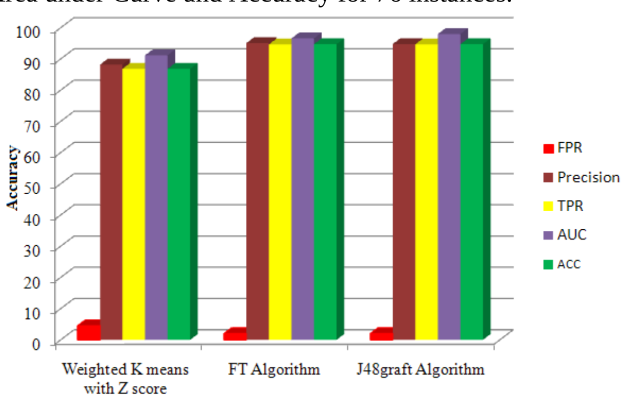


Fig. 8 Comparison of algorithms

The supervised algorithms FT and J48graft have accuracy 94.73%.

Different performance evaluation metrics can be applied to above algorithms. Weighted K-means with Z-score ranking is poor against classification algorithm as it has less TPR, Precision, AUC, ACC and more FPR. FT and J48graft have same TPR, FPR and ACC even then J48graft is best because it has higher AUC. Comparison of various evaluation metrics is shown in table 1.

TABLE I
COMPARISON OF ALGORITHMS BASED ON EVALUATION METRICS

Algorithm	FPR	Precision	TPR	AUC	ACC
Weighted K means with Z score	0.04	0.882	0.869	0.912	86.84
FT	0.02	0.951	0.947	0.966	94.73
J48graft	0.02	0.948	0.947	0.98	94.73

V. CONCLUSION

In the paper, performance of various clustering and classification algorithms is compared. For the study of brain tumor detection different variants of K-means could able to achieve accuracy between 50-86%. This much accuracy is not sufficient for medical application. Decision tree algorithms have performed better and quiet reliable for medical use. So, Author could conclude that supervised learning algorithms are more reliable than unsupervised learning algorithms in brain tumor detection. Author achieved 94.73% accuracy and 0.98 AUC using J48graft algorithm.

ACKNOWLEDGMENT

We wish to express our deep sense of gratitude to our guide Prof. M. B. Nagori for her valuable and firm suggestions, guidance, encouragement and constant support throughout this work without which it would not be possible to do this work. She took deep interest in checking the minute details of the paper and guided us throughout the same. She has been constant source of inspiration.

We also feel a deep sense of gratitude to Prof. V. P. Kshirsagar, HOD of Computer Science & Engineering department for his continuous encouragement, guidance and cooperation. We are also sincerely thankful to Dr. P. S. Adwani**,** Principal, Government College of Engineering, Aurangabad for being a source of motivation for this work

REFERENCES

[1] Magnetic Resonance, Functional (fMRI) – Brain, Retrieved from *www.radiologyinfo.org/en/pdf/fmribrain.pdf*, [2013-04-10].

[2] Jascap, "Brain Tumours", Retrieved from *http://www.healthlibrary.com/book77_chapter947.htm*

[3] Dina Aboul Dahab, Samy S. A. Ghoniemy and Gamal M. Selim, "Automated Brain Tumor Detection and Identification Using Image Processing and Probabilistic Neural Network Techniques", *International Journal of Image Processing and Visual Communication*, Vol. 1, Issue 2, pp.1-8, October 2012.

[4] Central Brain Tumor Registry of the United States (CBTRUS). Fact Sheet. Available at: http://www.cbtrus.org. Accessed April 4, 2013.

[5] American Cancer Society: Cancer Facts and Figures 2013. Atlanta, Ga: American Cancer Society, 2013. Available online. Last accessed April 6, 2013.

[6] P.Rajendran and M.Madheswaran, "An Improved Image Mining Technique For Brain Tumour Classification Using Efficient

classifier", *International Journal of Computer Science and Information Security*, Vol. 6, No. 3, pp. 107-116, 2009.

[7] hasan aydin, nilay aydin oktay, serdar spaholu, elif altin and baki hekmolu, "The Efficacy and Value of Proton MR Spectroscopy In Evaluating the Brain Tumours", *The New Journal of Medicine 2010*, pp. 37-42, 2010.

[8] T. Logeswari and M. Karnan, "An improved implementation of brain tumor detection using segmentation based on soft computing", *Journal of Cancer Research and Experimental Oncology*, Vol. 2, No. 1 pp. 006-014, March, 2010.

[9] G Vijay Kumar and Dr GV Raju, "Biological Early Brain Cancer Detection Using Artificial Neural Network", *International Journal on Computer Science and Engineering*, Vol. 02, No. 08, pp. 2721-2725, 2010.

[10] Andac Hamamci, Nadir Kucuk, Kutlay Karaman, Kayihan Engin, and Gozde Unal, "Tumor-Cut: Segmentation of Brain Tumors on Contrast Enhanced MR Images for Radiosurgery Applications", *IEEE TRANSACTIONS ON MEDICAL IMAGING*, Vol. 31, No. 3, pp. 790-804, March 2012.

[11] Sudipta Roy and Samir K. Bandyopadhyay, "Detection and Quantification of Brain Tumor from MRI of Brain and it's Symmetric Analysis", *International Journal of Information and Communication Technology Research*, Volume 2 No. 6, pp. 477-483, June 2012.

[12] Kathiresan V and Dr P Sumathi, "An Efficient Clustering Algorithm based on Z-Score Ranking method", *International Conference on Computer Communication and Informatics* Jan 2012.

[13] J. GAMA, "Functional Trees", *Machine Learning*, 55, pp. 219–250, 2004.

[14] Geoffrey I. Webb, "Decision Tree Grafting", *In Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence, Nagoya*, pp. 846-851, 1997.

[15] Geoffrey I. Webb, "Decision Tree Grafting From the All-Test-But-One Partition", *Machine Learning, IJCAI '99 Proceedings of the Sixteenth international joint conference on Artificial intelligence*, vol. 2, pp. 702-707, 1999.

[16] University of Waikato. *Weka 3: Data Mining Software in Java*. http://www.cs.waikato.ac.nz/ml/weka/index.html [2013-04-09]

BIOGRAPHY

Meghana B. Nagori has been working as an Assistant Professor in Computer Science and Engineering department since last 10 years. Currently author is carrying out research work in the area of developing models and methods for efficient detection and classification of Brain tumour.

Shivaji U. Mutkule received his B.E. degree in Information Technology from Vidya Pratishthan's College of Engineering, Baramati affiliated to University of Pune, India in 2011 and pursuing M.E. degree in computer science and engineering from Government Engineering College, Aurangabad, India. He has accepted the offer letter from Great Software Laboratory Pvt. Ltd Pune for the post of Software Engineer.

Praful B. Sonarkar received his B.E. degree in Computer Technology from Yashwantrao Chavan College of Engineering, Nagpur, India, in 2009 and pursuing M.E. degree in computer science and engineering from Government College of Engineering Aurangabad, India. His research interests include Digital Image Processing, Data Mining, Network Security, Artificial Intelligent, Machine learning and Bioinformatics.