International Journal of Computer Science Engineering Techniques (IJCSE) Volume 10 Issue 1 ,January 2026

ISSN : 2455-135X

Automated Brain Tumor Detection through Machine Learning Approaches

Ashima Gupta[1], Dr. Vivek Chaplot[2]

Department of computer science, Bhupal Nobles' University, Udaipur

ashimagupta509@gmail.com

Abstract - *We propose a novel and automatic computational approach to precisely identify subjects with Brain Tumor from normal brains. As due to high rate of manual error in data acquisition and interpretation there is a need to develop efficient algorithm capable of identifying early biomarkers and exemplify brain disease. So, we have developed improved method which uses adaptive moving mapping with FK-means and 22 GLCM features which were extracted from covariance matrix. The results of comparison experiments on real DICOM images taken from MB hospital (Udaipur) and other database demonstrate the effectiveness of the proposed method. Besides the effectiveness of tissue classification and tumor extraction it also specifies advantage of multifeatures combination to the single-feature method. Moreover, early used techniques faced high MSE, low PSNR and high computational time. This paper presents comparison with existing techniques like Fuzzy K-means and self-organising mapping over validating parameters.*

Keywords - AMKFSOM, Brain Tumor, clustering, feature extraction, Magnetic Resonance Imaging.

1. Introduction

A tumor is an undisciplined growth of cell in brain area which is basically identified using different imaging techniques among which MRI is most common. It produces a complete brain image which faces some challenges such as visual evaluations, manual re orientations and other time consuming errors. In developed countries, the statistical data reveals that people dies because of inaccurate and advance stage predictions of tumor[1] .

The brain tumors can be divided into two forms benign and malignant. Here benign is non-cancerous with controlled growth and does not affect other cells. It can be easily identified, removed and hardly ever grow back. Another is malignant brain tumor which grows rapidly (contains cancerous cells) and requires early diagnosis. refers to the procedure of partitioning a surrounding healthy brain tissue[3,4]. Many times imaging techniques does not locate objects and boundaries in images efficiently misidentification of tumor positions and misinterpretations . Medical images acquired by MRI produces visual representations of the interior of the brain which suffers certain issues such as noise , contrast, homogeneity and many more which could be critical issues for clinical analysis and medical intervention. So, efficient algorithm has to be designed which would acts as fully automatic system for analysing MRI images for diagnosis purpose. *Somasundaram et.al.* [8] implemented an automated brain extraction algorithm for axial T2 – W images. Mohammad et al.[10]used SVM classifier for tumor detection .Many used optimisation techniques for feature reduction [9].

In this paper ,we have used mapping method as it is efficient than prevailing methods along with FKM and GLCM .So we compared proposed FCM, FKM,SOM[7,9] and proposed on the basis of accuracy, mean square error, peak signal to noise ratio, Tanimoto index and Dice index which is defined in later sections. The comparison among various mapping method with their parameter is highlighted. All calculations were implemented using MATLAB. In later sections, its concepts and processing steps are explained.

2. Material and Methods

2.1 Self Organizing Map(SOM)

SOM (self-organised mapping) is one of the most common and efficient unsupervised learning neural network techniques. It gathers similar type of feature vector into two or more dimensional lattice which is mapped over output space. It is based on number of neurons and distance vector which expresses the lattice structure and weight factor. The training data corresponds to distance between input vector and corresponding weight factor with respect to output vector and neurons corresponds to the mapping area from high dimensional space onto a plane. There is a very important parameter to be calculated that is Best Matching Unit (BHU) whose value is updated[12,13].

The layers are typically located on a regular low-dimensional grid which may be either rectangular or hexagonal. A model of SOM network with M=10 inputs and L=10 x 10=100 neurons and data set with 300 inputs vector space is shown in figure 2 .Here input vector is $x(t) = [x_1,.., x_M]$ and each input is connected to all of neurons via corresponding weights $wj=[w_1,w_2…..w_N]$ where t=1,2,…,L ,i= 1 to Land j=1,2,…,M.

2.2 Adaptive Moving Fuzzy K-means Self Organizing Map(AMFKSOM)

Proposed algorithm(AMFKSOM) includes Proposed algorithm(AMFKSOM) includes integration of two algorithm AMSOM and FKM. The proposed method follows following steps[14,15]:

1. Initialize rectangular grid structure and size with number of neurons N .
2. Initialize weight vectors (wp) to random value based on number of features similar to SOM batch algorithm where $w_v(s + 1)$ is weight function at epoch (s+1), s is current epoch and s +1 is the next , $n_u(s)$ the number of patterns that are assigned to neuron u, Θ(u,v,s) is the neighborhood function and is a measure of how close are neuron u and neuron v, $x_u(s)$ is the mean feature vector of all x that are assigned to neuron u at epoch s, here ru ,rv are the vectors in the output space for neurons u and σ(t) is the adaptation factor, decreasing through training. So

$$p = nu(s).\Theta(u,v,s) \tag{1}$$

$$wv(s+1)=\frac{\sum p.xu(s)}{\sum nu(s).\Theta\ (u,v,s)} \tag{2}$$

$$\Theta(u,v,s)=\exp\left(-\frac{\|ru-rv\|^2}{2\sigma(s)^2}\right) \tag{3}$$

Lastly, the neuron vectors (wu) are computed when the neuron weight vectors update is over. These distances depict the closeness of neurons in the input space and output space which measure is used to update neuron positions[5].

$$rv(s+1)=rv(s)+\alpha(s)\frac{\sum nu(s).\delta uv(s).(ru(s)-rv(s)}{\sum nu(s).\delta uv(s)} \tag{4}$$

$$\Theta(u,v,s) = \exp\left(-\frac{\|wu-wv\|^2}{\gamma^*\sigma(s)^2}\right) \tag{5}$$

3. Initialize position vectors ($remove_{p1}$) as per initial grid structure (8 or 10). They are updated using two new neurons (p1 and p2) are calculated

$$k_{p1} =(1+\beta)^*k_p \tag{6}$$

$$k_{p2} =-(\beta)^*k_p \tag{7}$$

$$remove_{p1}=remove_{p1} \tag{8}$$

$$\text{remove}_{p1} = \frac{2*\text{remove}_{p1u} + 2*\text{remove}_{p}}{4} \tag{9}$$

4. Initialize edge connectivity matrix (Eg) values as per grid connectivity.
5. Initialize edge age matrix (Ag) values to zero.
6. Compute moving threshold (MT) as per dimension of data produced using GLCM matrix and then compute spreading factor (SF).It is given by

$$T= - \ln (D) *(-0.51) \tag{10}$$

7. In training phase find winner neuron N and increase such that neuron N is winner by 1.
9. Find second best matching neuron N and excluding Nb
10. Age all edges between N and its neighbors and increased by one. The epoch for each incident edge between neurons (i, j), if A(i, j) ≥agemax, then this edge is removed.
11. Connect N with Nb
12. Reset edge between N and Nb to zero
13. Update neuron weights and positions .
14. Update neuron position
15. If neurons need to be added/removed and update accordingly
16. If error does not change significantly then end step 7 else Continue
17. Smoothing phase involves fine-tune weights and deliver the AMSOM neuron weight vectors .
18. Feed the centroid to compute clustering using K means using

$$\text{Clust}=\sum_{i=0}^{k}\sum_{i=0}^{n} uf_{ij} *\Theta(u,v,s) \tag{11}$$

2.3 Methodology

In our proposed system we get the benefits of two algorithm in which one helps in fast clustering the region and other efficiently classifies using mapping. As shown in Figure 1 it consists of five stages : pre-processing ,clustering stage, feature extraction , classification and validation stage. The main idea of proposed technique is to reduce the number of iterations and dimension reduction of data[2]. Which in further reduces computational time and increases efficiency. Feature extraction identifies some specific properties from original datasets which generates distinguishing patterns. Thus, extracted features are self-revealing and distinct which acts as important key factor or biomarker for classification. The feature vector of the MRI image is calculated by I order and II order features which is produced using covariance matrix. First order features and second features are intensity, average intensity, variance, hu moment, entropy, energy, homogeneity, angular moment and correlation. The features are extracted using gray-level co-occurrence matrix (GLCM) from MRI image of 256x256 size. The GLCM is a pixel with gray-level (grayscale intensity) value i occurs horizontally adjacent to a pixel with the value j. Each vector p(i,j) in GLCM specifies the number of times that the pixel with value i occurred horizontally adjacent to a pixel with value j. Here I order feature means pixel independent of relationship factor which includes covariance, intensity and mean where as II order includes distinct tissue factors such as hue factor and features dependent on orientations. As second order features, are invariant under rotation and scaling. List of features used in Table 1are:Mean,standard deviation(Std),median, autocorrelation(ACr),clustercrominence(CP),clustershade(CS),contrast(C),correlation(Cr),differententropy(DE),differentvariance(DV),Dissimilarity(DS),energy(Eg),entropy(En),Homogenity(Ho),information measure of correlation1(IMC1),information measureofcorrelation2(IMC2),Inversedifference(ID),Maximum probability(MP),Sum average(SA),Sum entropy (SE),Sum of squares(SS),Sum variance(SV).

Preprocessing is implemented by using series of pre-processing steps on MRI image is done before any special purpose processing. This involves skull removing and denoising which removes corrupted noise such as Gaussian and poison noise. Then comes clustering in which set of feature vectors are fed to proposed clustering algorithm which is developed using FKM and AMSOM .AMSOM is used for initial clustering and dimensionality reduction of MRI image. Here real brain image is acquired from hospital and reshaped to 256 x256 pixel size which is reduced through AMSOM. Initially random vectors are chosen for weight factors then winning neurons with respect to weight factors are calculated. The Euclidean distance is calculated at each point help to map output image to a corresponding point. The updated weight function is resultant value obtained using AMSOM and the weight function is fed to FKM algorithm. So the limitation of FKM of handling large and huge data is easily reduced with inclusion of AMSOM .The Figure 1 shows the block diagram of the methodology carried out for proposed method,

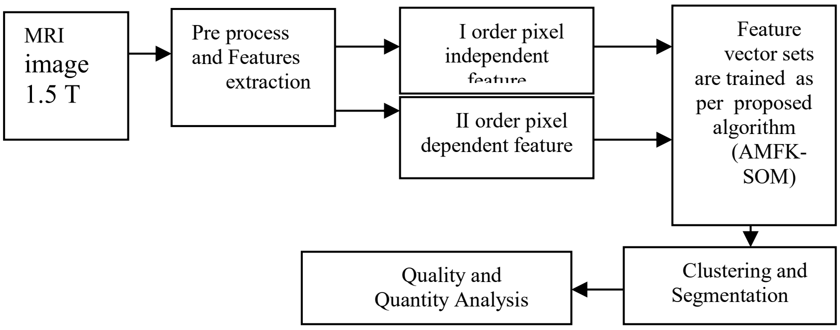


Fig.2 . Block diagram of the Proposed method

In the validation stage ,the segmented images are compared with ground truth as illustrated in experimental results using following equations.

These are unitless quantity but can be expressed in percentage if multiplied by 100.

Mean Square Error (MS) is defined as squared error [14,16]value between the input image M(i,j) and the segmented image S(i,j).It is given by

$$MS=\frac{1}{m}\sum_{i=0}^{m=1}\sum_{j=0}^{n=1}[M(i,j)-S(i,j)][M(i,j)-S(i,j)] \qquad (15)$$

where ‘m’ & ‘n’ denotes the matrix size in an input image.

Peak Signal to Noise Ratio(PS) is defined as opposition of an image to noise signals and if high, then the impact of noise interference to MR brain image is quite low. The PS values less sensitive to noise signals lies between 40 to 100 dB. PSNR is given by:

$$PSNR=10\log\frac{\text{optimal value}^2}{MS} \qquad (16)$$

Tanimoto index (TC) is defined as the ratio of the intersection between the input image (M) and the output image (S) to the union function of the input image (M) and the segmented output image (S) .

$$TC=\frac{(M)}{(M\cup S)}\cap\frac{(S)}{(M\cup S)} \qquad (17)$$

Dice Overlap Index (DOI) is defined by overlapping function of the input image (M) and the clustered output image (C) . The Table 1 shows the extracted features for input images in which features are around 22.

$$DOI=\frac{(4*TC)}{2(1+TC)} \qquad (18)$$

Table 1 Feature extracted dataset

No.	Features	Image1	Image 2	Image3	Image4	Image5
1	Mean	0.2234	0.2035	0.2676	0.2046	0.2057
2	Std	0.0808	0.0748	0.0788	0.0646	0.0957
3	Median	0.2716	0.2324	0.4422	0.3294	0.3608
4	ACr	8.8084	7.689	10.90	7.6537	8.3197
5	CP	234.6254	99.6252	165.42	91.5626	195.7736
6	CS	21.3709	7.5505	-0.5218	5.47	17.056
7	C	0.1663	0.0528	0.0781	0.108	0.1463
8	Cr	0.9686	0.9879	0.9860	0.974	0.9713
9	DE	0.4170	0.1999	0.2676	0.3162	0.3746
10	DV	0.1476	0.0504	0.0740	0.0998	0.1326
11	DS	0.1368	0.0496	0.0740	0.0923	0.1171
12	Eg	0.2380	0.3982	0.2792	0.3240	0.3299
13	En	1.9926	1.2442	1.6046	1.4956	1.6273
14	Ho	0.9346	0.9754	0.9634	0.9554	0.9443
15	IMC1	-0.7165	-0.8304	-0.8035	-0.7515	-0.7249
16	IMC2	0.9444	0.9105	0.9403	0.9137	0.9180
17	ID	0.9361	0.9756	0.9066	0.9536	0.9458
18	MP	0.4380	0.5194	0.4066	0.4802	0.5105
19	SA	4.9982	4.7019	5.711	4.7642	4.8330
20	SE	1.8526	1.2030	1.5457	1.4131	1.5148
21	SS	2.6461	2.1891	2.7931	2.033	2.5534
22	SV	10.4181	8.7037	11.09	8.0251	10.067

3. Results and Validations

The experimental setup includes HP laptop with Intel(R) Core(TM) i3-5005U CPU @ 2.00GHz processor with 4 GB RAM using software version MATLAB (R2013a).The results of our proposed algorithm is obtained from 12 real datasets of MRI T1 images of different hospitals. Here, we have shown 5 distinct images of dataset I which is acquired from Dr. K.G. Srinivan , MD,RD ,Consultant Radiologist and Dr. Usha Nandini ,DNB, KGS Advanced MR & CT Scan -Maurai,Tamilnadu, India through Govindaraj Vishnuvarthanan and dataset II is acquired from Radiologists Dr. Ranbir Singh and Consultant Saji P.Mathew from Maharana Bhopals Hospital ,Udaipur. Proposed algorithm with validating parameters are calculated in Table 2 illustrating how are algorithm is better than others over ground truth data.. The validating parameters are calculated which shows low mean square value with respect to previous algorithm such as FCM,FKM and SOM-FKM . Here PSNR is improved as high value shows good and effective clustering.

Table 2 Parameters calculated using AMKSOM

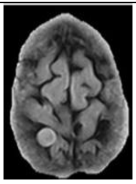
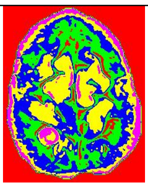
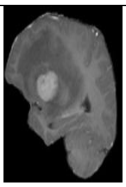
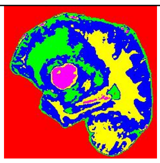
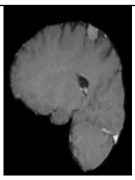
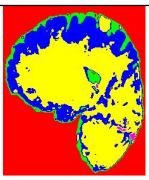
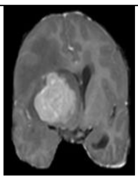
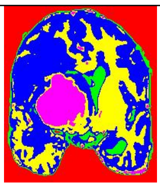
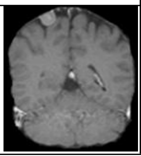
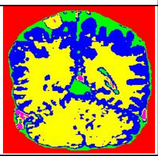
Parameters	Image 1	Image 2	Image 3	Image 4	Image 5
MSE	0.1	0.8	0.12	0.08	0.09
PSNR(dB)	58.14	58.66	57.17	58.91	58.48
DOI	0.02	0.197	0.0183	0.02	0.0197

As shown in Table 3 ,we used same number of clusters for FCM,FKM,SOM-FKM and proposed algorithm to evaluate under same conditions .It shows that with K=3 FCM gave good results but with 59 iterations not accurately ,whereas in FKM with iteration number 49 results were good with 5 clusters .SOM-FKM gave better results compared to Fuzzy means(FCM and FKM) . The Table 4 shows the clustered image with extracted tumor on the basis of 22 features extracted. Here output images are produced using AMKSOM clustering where (a) – (c) are input and images of dataset 1and (f) – (g) are input and output images of dataset 2.

Table 3 Performance parameters for different algorithms

Parameters	FCM	FKM	SOM-FKM	Proposed
MSE	2.880	1.9	0.15	0.01
PSNR (dB)	43.57	45.17	56.26	68.16
Accuracy (%)	85	82.5	82.75	89.11
Similarity Criteria (%)	89.6	91.0	92.0	96.8

Table 4 Segmented image using AMKSOM with tumor extraction

Dataset I				Dataset II	
S. No.	**Input image**	**Clustered image**	**S. No.**	**Input Image**	**Clustered image**
(a)			(f)		
(b)			(g)		
(c)					

We found that the processing time of images with skull structure is more in comparison to skull free images. The performance matrix is explained below. The results of each technique are recorded in the following table according to execution time, sensitivity and performance parameters.

The Figure 3 illustrates the improvement in computational processing time. Our proposed algorithm has reached up to 16 seconds which is and histogram presentation of precision rate illustrating the efficiency of proposed clustering algorithm for 5 set of images is shown in figure 4.

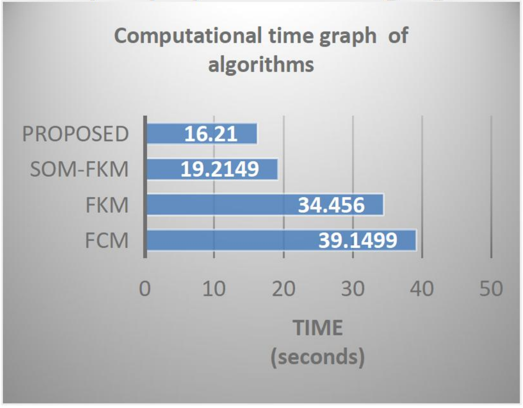


Figure 4 Graphical analysis of processing time

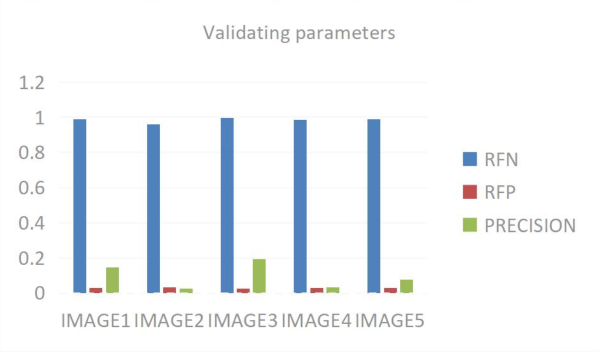


Figure 5 Graphical analysis of validating parameters

4. Conclusion

The MRI image analysis plays an important role disease diagnosis. With advent of new technology there are wide range of techniques such as MRI, CT scan, PET scan and many more. The MRI is better than CT scan but expensive so computational processing may reduce the cost of patient's expense if diagnosed at early stage. The K- means can detect faster than FCM but fails to cluster image data with noise. SOM-FKM has improved dimensional reductionality but fails to detect with huge data .So, the new novel attempt to validate the improved mapping algorithm AMKSOM for disease detection is carried out in this paper. From the experimental results, we proved the effectiveness of AMKSOM which is efficient enough in satisfying are goal needs. The manual error made by physician or any other leads to delay of treatment and ignorance which can be avoided with our algorithm. This helps the patient to get treatment in the earlier stage of the tumor to avoid severity. The result produced using proposed algorithm reveals satisfactory. It can be further extended by implementing on PET images and 3-D imaging techniques for other diseases.

References

1. Archya Dasgupta, Tejpal Gupta, and Rakesh Jalali(2016).Indian data on central nervous tumors: A summary of published work, South Asian J Cancer,Jul-Sep;ed 5 vol. 3:pp.147–153.doi: 10.4103/2278-330X.187589[accessed:28/3/2018]
2. Govindaraj V, Vishnuvarthanan A, Thiagarajan A, Kannan M, Murugan PR(2016). Short Notes on Unsupervised Learning Method with Clustering Approach for Tumor Identification and Tissue Segmentation in Magnetic Resonance Brain Images, J Clin Exp Neuroimmunol vol 1, pp. 101,2016.doi:10.4172/jceni.100010.
3. N.NandhaGopal (2013). Automatic Detection of Brain Tumor through Magnetic Resonance Image. International Journal of Advanced Research in Computer and Communication Engineering 2(4).
4. El-Hachemi Guerrout, Ramdane Mahiou and Samy Ait-Aoudia(2013).Medical Image Segmentation on a Cluster of PCs using Markov Random Fields, International Journal of New Computer Architectures and their Applications (IJNCAA). The Society of Digital Information and Wireless Communications (SDIWC) 3(1): 35-44 ,ISSN: 2220-9085.
5. Ortiz , J.M. Górriz , J. Ramírez , D. Salas-González , J.M. Llamas-Elvira(2013) .Two fully-unsupervised methods for MR brain image segmentation using SOM-based strategies", Applied Soft Computing, Science Direct13, 2668–2682.
6. Sultan Aljahdali and E. A. Zanaty (2012).Automatic Fuzzy Algorithms for Reliable Image Segmentation. International Journal for Computers & Their Applications, 19(3).
7. Ashraf Anwar and Arsalan Iqbal(2013).Image Processing Technique for Brain Abnormality Detection, International Journal of Image Processing 7(709).
8. Somasundaram K, Kalaiselvi T(2012).Fully automatic brain extraction algorithm for axial T2-weighted magnetic resonance images. PubMed Comput Biol Med 40: 811-822.
9. P.Vasuda and S.Satheesh(2010).Improved Fuzzy C-Means Algorithm for MR Brain Image Segmentation. (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 05, 1713-1715.
10. Mohammad Javad Abdi, Seyed Mohammad Hosseini, and Mansoor Rezghi(2012).A Novel Weighted Support Vector Machine Based on Particle Swarm Optimization for Gene Selection and Tumor Classification ,Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine , Article ID 320698, 7 . doi:10.1155/2012/320698.
11. Eman Abdel Maksoud,Moh. Emlogy,Rashid Al-Awadi(2015).Brain tumor segmentation based on hybrid clustering technique. Egyptian Informatics Journal (2015) 16, 71–81.
12. Kruti Choksi, Bhavin Shah, Ompriya Kal(2016).Intrusion Detection System using Self Organizing Map: A Survey, Int. Journal of Engineering Research and Applications ISSN : 2248-9622,Vol. 4, Issue 12(Part 4), pp 11-16.
13. Nameirakpam Dhanachandra, Khumanthem Manglem and Yambem Jina Chanu(2015).Image Segmentation using K-means Clustering Algorithm and Subtractive Clustering

Algorithm ,Procedia Computer Science 54 , 764 – 771, Published by Elsevier B.V. under the Eleventh International Multi-Conference on Information Processing-2015 .

14.Spanakis, Gerasimos, and Gerhard Weiss(2016).AMSOM: Adaptive Moving Self-organizing Map for Clustering and Visualization, arXiv preprint arXiv:1605.06047(2016) .

15. Emal Abdel Maksoud,Moh. Emlogy,Rashid Al-Awad(2015).Brain tumor segmentation based on hybrid clustering technique. Egyptian Informatics Journal,Science Direct,16,71-81.

16.Sudipta Roy, Debnath Bhattacharyya, Samir Kumar Bandyopadhyay & Tai-Hoon Kim (2017). An Iterative Implementation of Level Set for Precise Segmentation of Brain Tissues and Abnormality Detection from MR Images, IETE Journal of Research, DOI: 10.1080/03772063.2017.1331757.