

Brain Tumor Detection & Features Extraction From MR Images Using Segmentation, Image Optimization & Classification Techniques

Zahoor Ahmad¹,

Deptt. of Computer Science & Information Technology,
University of Engineering & Technology
Peshawar, Pakistan

Engr. Seemab Gul²,

Deptt. of Electrical Engineering,
University of Engineering & Technology
Peshawar, Pakistan

Abstract—Nervous system most vital part is brain. Our actions whether voluntary or involuntary all depend on it. The health of the brain is very crucial. A number of deaths are caused by brain related diseases. Brain Tumors are one of the dangerous diseases that damages the brain and is sometimes incurable. Tumors are formed when brain cells split up in an abnormal fashion for various reasons.

Timely identification of brain tumors is very crucial for curing this disease. Many techniques are developed by researchers for diagnosis of brain tumors with the help of X-Ray, CT Scan and Magnetic Resonance Imaging (MRI), however MRI scan is considered to be an optimal solution for brain tumor identification due to its harmless effects on human body. Examination of brain tumor tissues is very complex and without clear detailed study by radiologist can lead to loss of precious human life. Moreover manual diagnosis of MR images is more prone to human errors. Due to these issues – computer aided diagnosis using efficient algorithms are favoured over manual diagnosis of brain tumors.

In biomedical engineering, various methods have been employed for detection, segmentation and classification of brain tumor, but they contain various shortcomings in one form or other. This paper is focused on evaluating the dataset of brain MRIs and are enhanced by employing various image processing and segmentation techniques before classifying the MR image in tumor and non-tumor images. Furthermore – performance of popular classification techniques such as SVM, KNN and ANN are evaluated by training and testing them with same set of MR images. The output results are computed by statistical analysis.

Keywords—MRI, GLCM, Watershed Segmentation, SVM, KNN, ANN, Brain Tumor

1. INTRODUCTION

In biomedical sciences, the conventional technique for tumor detection, identification and classification is through manual human inspection. The manual inspection method is more prone to human error, time consuming and at times very impractical solution for analysing huge quantity of MR images. Diagnosing MR images is very laborious and require keen attention and accuracy as wrong diagnosis could lead to loss of precious human life.

Human brain is the most complex part of human body. Brain controls many complex functions. Latest development in detection process of brain tumor using MR (Magnetic

Resonance) images and medical imaging is focusing more on real time observation and analysis of tumors by making use of more stable and reliable algorithms. Moreover, images obtained from Computed Tomography (CT) scan have become an active and operating area of research so far. However MR Images are favored over CT images since it don't produce any harmful radiation. One out of the most issuing problems for most of imagery diagnosis systems in medical is the partitioning of cells and the nuclei of these cells, from the rest of content in image.

With advancement in technology and development of artificial intelligent techniques the aforementioned issued could be overcome. The implementation of such algorithm is much easier and its outcome are superb. The proposed work focuses on correctly identifying brain tumors and non-tumor MRIs based on segmentation techniques followed by machine learning and classification. The MR images would be enhanced through image processing techniques by applying different filters and converting to grey scale. Furthermore, after image segmentation features would be extracted through GLCM, afterwards through the extracted features different supervised classification techniques such as SVM, KNN, and ANN would be trained and tested for classification of tumor and non-tumor MR images. The performance of proposed methodology is evaluated by statistical analysis and confusion matrix.

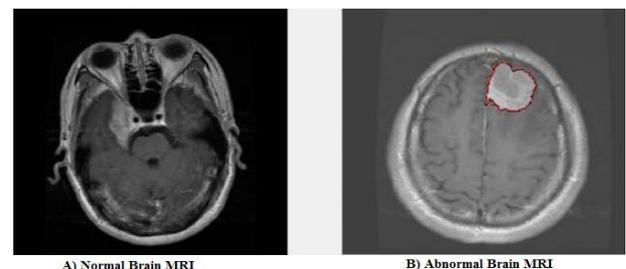


Fig. 1: Normal & Abnormal Brain MRIs

2. EXISTING WORK

By analyzing the literature it could be concluded that different studies have been carried out using various techniques for segmentation, classification and identification of tumor in MR images but they have various pros and cons.

M. Masroor Ahmed et al [1] proposed the method of the brain tumor detection using Kmeans Clustering. Nagalkar VJ et al [2] proposed brain tumor detection using soft computing method. This method can cause false detection in seeing scan. Rajesh C. Patil et al [3] proposed the method of the brain tumor extraction from MRI images using MATLAB.

Shivakumarswamy, Akshay Patil, Chethan, Prajwal, Sagar.V.Hande [4] have focused on KNN classification in their research. It is a multistep technique whose first step is acquiring an MRI scanned image. The image is pre-processed for noise removal and size changings. Then it is forwarded to segmentation by two algorithms: K-Means clustering and Fuzzy C-means segmentation. Then the tumor cells are separated from normal cells and the area of the diseased part containing tumor is measured. The stage of the tumor is determined according to this calculated area. Finally, the concerned people are informed about the determined stage of the tumor.

Saha, Ray, N., B. N. Greiner, R., Murtha, A., & Zhang, H [5] have explained the outcomes in their paper. The study of patients comprises of MR slices, is the input and a subset comprises of the slices in which including axis-parallel boxes which circumscribe the tumors, is the output. Their proposed approach is basically an unsupervised method for detection of change which searches the dissimilar region (dissimilar region are the bounding boxes axis-parallel) between the two halves of brain, one left and one right in MR slice's axial view.

Dunn [6] uses fuzzy c-mean (FCM) clustering algorithm for image segmentation. In clustering most important task is finding cluster to classify pixels [7]. In this manner, fuzzy method is fast and accurate.

P. Mohanaiah, P. Sathyanarayana, L. GuruKumar [8] applied neuro-fuzzy segmentation process on the data attained from MRI images is to locate several tissues like white matter, cerebrospinal fluid, grey matter and tumor. The benefit of fuzzy means algorithms and hierarchical self-organizing map help in categorization of the image layer by layer. The lowest level weight vector is attained by the abstraction level. The researchers have examined tumor segmentation based on edge detection algorithms and watershed in HSV color model.

Bhagwat et al [9] they showed K-means algorithm produces more accurate result compared to Fuzzy c-means and hierarchical clustering. Most of the algorithms in this field are developed by inspiring or improving k-means [10]. The algorithm upgrades the clusters iteratively and runs in a loop until it reaches to optimal solution. But Performance of K-means algorithm depends on initial values of cluster centers [11]. Therefore the algorithm should be tested for different outcomes with different initial cluster centers by multi-running.

Another technique has been discussed in literature that is based on thresh-holding for detection of tumors. Kalaiselvi, T. & Sriramakrishnan [12] have used this approach in their paper. They suggested a model for brain tumor identification

from the magnetic resonance imaging (MRI) of human head scans. The MRI images are pre-processed by transformation techniques and thus enhance the tumor region. Then the images are checked for abnormality using fuzzy symmetric measure (FSM). If abnormal, then Otsu's thresh-holding is used to extract the tumor region.

V.Vani, M. Kalaiselvi Geetha [13] used SVM, KNN and Decision Tree (DT) for classification of tumors and automating process of brain tumor detection. With decision tree they claimed to achieve to 98% accuracy.

Nikita V. Chavan, B.D. Jadhav, P.M. Patil [14] proposed 02 steps methods for detection and classification of brain tumors. In their proposed methodology they have used Grey Level Co-occurrence method (GLCM) for extracting tumor features and in second stage they have used K-NN classifier. They have used supervised machine learning algorithm for classification of benign stage tumor. With KNN they have achieved 96% accuracy.

2.1 Analysis of Existing Methods

Fig. 2 depicts the various methods developed and used for analyzing brain related abnormalities. These studies proved pivotal and based on these innovations are made to existing methods for achieving more accurate results. The shortcomings against these studies are presented in the table 1.

TABLE 1: Analysis of Existing Methods

Paper	Authors	Year	Techniques	Results/Limitation
An Artificial Neural Network Approach for Brain Tumor Detection Using Digital Image Segmentation	Kamal Kant Hiran, Ruchi Doshi	2013	Segmentation using Neural Network	Brain tumor segmentation explained while classification is not elaborated
Image Processing Techniques for Brain Tumor Detection	Vipin Y. Borole , Sunil S. Nimbhore , Dr. Seema S. Kawthekar	2015	Classification	When only classification is applied, it ignores the poor quality images.
Brain Tumor Detection Using MRI Images	Pranita Balaji Kanade , Prof. P.P. Gumaste	2015	SVM	SVM is not very scalable in dealing with large number (billions) of training data.
Brain Tumor Segmentation and Extraction of MR Images Based on Improved Watershed Transform	J.Mehena, M. C. Adhikary	2015	Watershed	Segmentation of MRI brain tumors in both dimensions.
Brain tumor detection using Image processing and sending tumor information over GSM	Shivakumarswamy G.M. , Akshay Patil.V.	2016	K-Mean and Fuzzy C Mean	Results in distorted boundaries and edges
Brain Tumor	Rajeev Kumar ,	2017	Morphological	Not works for

Segmentation by Modified KMean with Morphological Operations	Dr. K. James Mathai		Operators and K-mean	global cluster
A Survey on Brain Tumor Detection Using Image Processing Techniques	Luxit Kapoor Sanjeev Thakur, Professor	2017	Tumor Detection using threshold operation in MRI Brain Images	The method requires estimating threshold and does not produce accurate results most of the time. It is expensive as well.
Brain Tumor Detection in MRI Images With New Multiple Thresholding	Sandeep Patel, Divyanshu Rao	2017	Brain Tumor Detection and Segmentation Using Histogram Thresholding	Useful for linear image but not for detecting Brain tumors. Doesn't give accurate results.
An Automatic Brain Tumor Detection, Segmentation and Classification Using MRI Image	Arbaz Mukaram Chidananda Murthy.M.V, M.Z.Kurian	2017	Classification	When only classification is applied, it ignores the poor quality images.
Efficient image segmentation of brain tumor detection using fuzzy c-mean and mean-shift	Mandip kaur, Prabhpreet kaur	2017	Fuzzy c-mean and mean-shift	Neglected the use of fuzzy and region growing segmentation

3. PROPOSED METHODOLOGY

Our proposed approach is aimed at achieving high accuracy in classifying MR image to Tumor and non-tumor target classes. Combination of supervised classification techniques such as SVM, KNN and ANN are trained and their performance is assessed by confusion matrix.

Our proposed methodology for detection and classification of brain tumors comprises of following steps. The main modules are pre-processing, morphological processing and then computing segmentation through watershed algorithm followed by feature extraction through Grey level co-occurrence matrix (GLCM), training and testing of classification techniques.

The data of MR images has been acquired from figshare which is uploaded by Jun Cheng for research purposes.

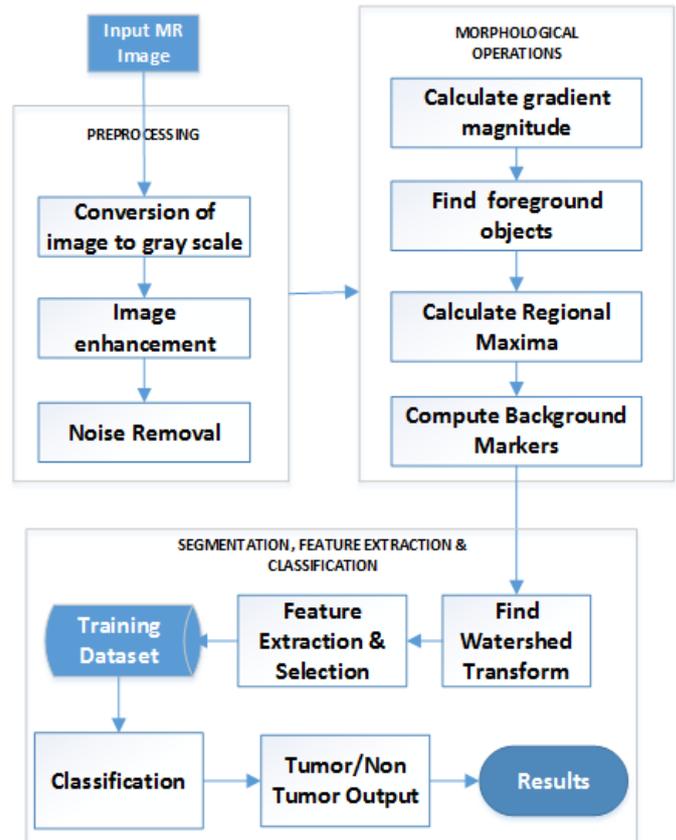


Fig. 3: Workflow of Proposed System

3.2. Image Pre-Processing & Noise Removal

The first step of our proposed method is image preprocessing. MR Image could be subjected to noise which can occur due to many factors such as electronic disturbance, low light and channel noise etc. As a result image could be blur at certain areas, for which image pre-processing is required. For noise removal and image enhancement it is first converted to grey scale, then gradient magnitude is calculated for segmentation by applying sobel edge mask, and Gaussian filter is applied for noise removal along with other arithmetic computation.

3.3. Applying Morphological Operations

In order to find foreground objects, morphological techniques called opening and closing by reconstruction are applied for cleaning up image. After applying these operations flat maxima inside each object is created which can easily be located.

3.4. Watershed based Segmentation

Segmentation technique is to filter out the area of interest. In our case it is the tumor region from MR image. Segmentation identifies boundaries on the basis of colour, texture and intensities. The segmentation can be done by edge detection method, by Region growing or by thresholding segmentation technique. For our work, we have employed region growing and watershed algorithm. Region growing mostly depends on the selection of threshold value by selecting group of seed pixels from original image [13], [14].

Watershed segmentation is a very helpful method that group up pixels on the basis of intensity from an image. Pixels having similar intensities are given a higher weight and grouped together.

3.5. Features Extraction using GLCM

Feature Extraction is the most critical phase in image processing. It is employed for reducing original dataset by measuring certain features. Features of the tumor and non-tumor MR images are extracted from the region separated from watershed segmentation using grey level co-occurrence matrix (GLCM) and are recorded separately in txt file. Features are an indicator for testing the presence of tumors.

3.6. Training & Testing of Algorithms

Features are fed into a classifier during training of data. After that an architecture is obtained which can classify the input image based on its learning. Once the features for tumor and non-tumor MR images are extracted and stored in txt files, they are then used to train different algorithm, in our proposed work since we want to evaluate the performance and accuracy of SVM, KNN and ANN so we trained and tested these algorithms with features of tumor and non-tumor MR images for MR classification.

3.7. MRI Classification

In machine learning classification is used for identification of new observation that is to which class this observation belongs. In our research we are using Support Vector Machine (SVM), K-nearest neighbour (KNN) and Artificial Neural Network (ANN) for classification of MR images into tumor and non-tumor classes.

3.8. Algorithm Steps for proposed System

Step.1: Input image in .mat format is read

Step.2: Input image is converted to grey scale. Strcat and cjdta function are used

Step.3: Gradient magnitude is computed and sobel edge mask along with Gaussian filter is applied for noise removal

$$w(T) = C(\sigma) \exp\left(-\frac{T^2}{2\sigma^2}\right); T = -1, 0, +1$$

Where,

T- is the distance in time from the current moment

σ - Standard deviation (sigma) is the parameter of the Gaussian filter, by default sigma is 0.5.

$C(\sigma)$ - is the normalization constant chosen to make the sum of all weights equal to the unit value.

Fig. 4: Image preprocessing using Gaussian filter

Step. 4: To compute foreground objects, morphological operations are applied and flat maxima is created

$$F \ominus B \subseteq F \subseteq F \oplus B$$

Where, F denote a gray scale brain image, B denote structuring element. Opening and closing operations satisfy:

$$F \circ B \subseteq F \subseteq F \cdot B$$

Fig. 5: Morphological Operations

Step. 5: Watershed segmentation is applied to get markers and object boundaries superimposed on original image

Step. 6: Through GLCM features of tumor & non-tumor MR images are extracted and stored in txt file.

Step.7: Training file for SVM and KNN is created from features file

Step.8: Training and testing of ANN is performed using matlab pattern recognition nprtool.

4. RESULTS AND DISCUSSION

In order to evaluate and validate the performance of SVM, KNN and ANN algorithms quantitatively the first step is to prepare the data for it. In our proposed work we are using brain tumor and non-tumor MR images dataset which is uploaded to figshare by Jun Cheng for research purposes. The dataset consisted of 3064 MR images.

The performance and accuracy of algorithm depends on quality data preparation, as the data is used for training and against training input data from tumor and non-tumor classes are given along with target classes for classification. The dataset in our case was huge and tremendous manual efforts were needed to separate tumor and non-tumor MR images.

For our work, we have analysed 2000 MR images and extracted 507 samples of MR images data. They have been stored separately in tumor and non-tumor files by manual examination. The rest were discarded mostly because of poor segmentation using watershed. In 507 samples of data 318 MR images are those containing no tumors while 189 MR images are those containing tumor images.

Features of 318 non-tumor categorized MR images are extracted through GLCM and stored as matrix in non-tumor txt file. The features of non-tumor MR images have been labelled as '0' while features of tumor MR images have been labelled as '1'. Target groups containing all labels of non-tumor and tumor MR images have been made and stored separately.

4.1. Classifiers Performance Evaluation by Confusion Matrix

Table 2 depicts generalized confusion matrix technique for brain tumor detection. Through confusion matrix better idea can be depicted regarding correct and incorrect instances of data as calculating accuracy alone could be misleading.

TABLE 2: Confusion Matrix for Brain Tumor Detection

N=Total Cases	Target Class Non-Tumor Output (0)	Target Class Tumor Output (1)
Non-Tumor Input (0)	TN	FP
Tumor Input (1)	FN	TP

System accuracy can be defined mathematically as follows:

$$\text{System Accuracy} = \frac{TP + TN}{(TP + TN + FP + FN)} \times 100 \tag{1}$$

If the count of True Positives (TP) and True Negatives (TN) are high then the system depicts high accuracy.

4.1.1. Sensitivity

When the system correctly classify tumor as tumor then it's sensitivity and true positive rate is calculated as follows:-

$$\text{Sensitivity} = \frac{TP}{(TP + FN)} \times 100 \tag{2}$$

4.1.2. Specificity

When the system correctly classify non-tumor as non-tumor then it's called specificity and true negative rate is calculated as follows:-

$$\text{Specificity} = \frac{TN}{(TN + FP)} \times 100 \tag{3}$$

4.2. Output of SVM Classifier

Since we want to evaluate to performance of SVM, so we trained SVM with our dataset. Table 3 lists the outcome of all input MRIs tests conducted through SVM technique. Overall 85% accuracy is achieved through SVM.

TABLE 3: Output of SVM Classifier

N=507	Target Class Non-Tumor Output (0)	Target Class Tumor Output (1)
Non-Tumor Input (0)	TN=272	FP=46
Tumor Input (1)	FN=30	TP=159

$$\text{System Accuracy} = \frac{159 + 272}{(159 + 272 + 46 + 30)} \times 100$$

System Accuracy = 85%

$$\text{Sensitivity} = \frac{159}{(159 + 30)} \times 100$$

Sensitivity = 84%

$$\text{Specificity} = \frac{272}{(272 + 46)} \times 100$$

Specificity = 85%

4.3. Output of KNN Classifier

Output of KNN classifier are listed in Table 4, KNN algorithm is trained on same dataset and each instance of output is recorded. Overall 88% accuracy is achieved through KNN.

TABLE 4: Output of KNN Classifier

N=507	Target Class Non-Tumor Output (0)	Target Class Tumor Output (1)
Non-Tumor Input (0)	TN=281	FP=37
Tumor Input (1)	FN=21	TP=168

$$\text{System Accuracy} = \frac{168 + 281}{(168 + 281 + 37 + 21)} \times 100$$

System Accuracy = 88%

$$\text{Sensitivity} = \frac{168}{(168 + 21)} \times 100$$

Sensitivity = 88%

$$\text{Specificity} = \frac{281}{(281 + 37)} \times 100$$

Specificity = 88%

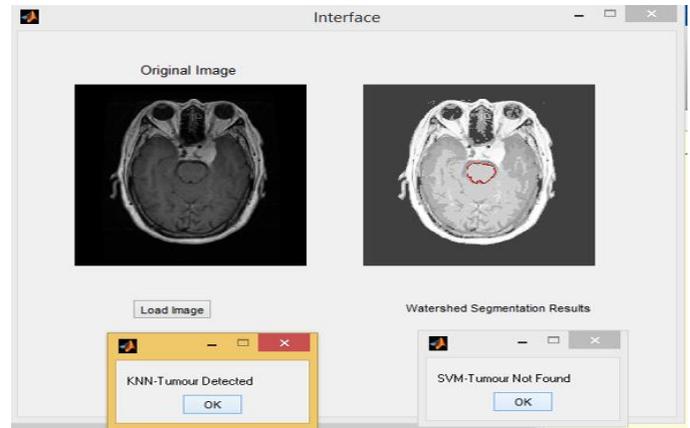


Fig. 6: Output of KNN & SVM Classifier

Figure 6 above depicts the output of our proposed system, where after training of KNN and SVM, input MR image containing tumor is given to the system. KNN technique correctly detected and classified tumor while SVM technique didn't detected tumor based on it's learning.

4.4. Output of ANN Classifier

For ANN classifier, neural network pattern recognition tool is used, where features of tumor and non-tumor stored in txt file is given as input to the network, while target groups are also specified. The data is then divided for training, validation and testing of network. 70% of the data has been used for training, while 15% each used for validation and testing of network. As shown in figure 7, the artificial neural network has been setup with 10 hidden neurons. The network has been trained using Bayesian regularization back-propagation function. Accuracy and errors of neural network varies with retraining of the network.

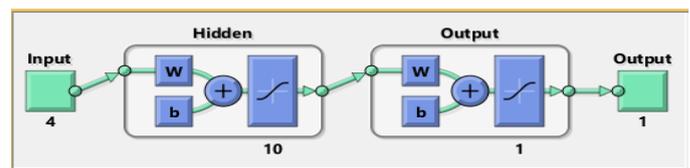


Fig. 7: Artificial Neural Network Setup

ANN confusion matrix is generated using matlab nprtool as shown in figure 8, out of 318 non-tumor MR images 298 are correctly classified as 0 or non-tumor target group while 18 non-tumor images are incorrectly classified as 1 or tumor class target group. In 189 tumor MR images 172 are correctly classified as tumor target classes while 17 are incorrectly classified as non-tumor target classes.

Confusion Matrix

Output Class	0	1	2
	0	1	2
	Target Class		
0	298 58.1%	20 3.9%	93.7% 6.3%
1	17 3.3%	172 33.9%	93.7% 6.3%
2	94.6% 5.4%	89.5% 10.5%	92.7% 7.3%

Fig. 8: Output of ANN Classifier

$$\text{System Accuracy} = \frac{172 + 298}{(172 + 298 + 20 + 17)} \times 100$$

System Accuracy = 92%

$$\text{Sensitivity} = \frac{172}{(172 + 17)} \times 100$$

Sensitivity = 91%

$$\text{Specificity} = \frac{298}{(298 + 20)} \times 100$$

Specificity = 93%

4.5. Cumulative Evaluation of SVM, KNN & ANN Classifiers

SVM, KNN and ANN performances are evaluated by calculating their system accuracy, sensitivity and specificity on 507 sample of MR data, the cumulative summary of which is given in below Table 4:-

TABLE 5: Cumulative Evaluation of SVM, KNN & ANN

Classifiers	Performance Measures		
	Sensitivity	Specificity	Overall Accuracy
SVM	84%	85%	85%
KNN	88%	88%	88%
ANN	91%	93%	92%

As evident from table 5, ANN classifier has highest ratio of sensitivity, specificity and overall accuracy, which can further be increased by adjusting neurons, re-training of the network and by increasing data.

5. CONCLUSION & FUTURE WORK

In this paper, a comprehensive framework for segmenting tumor and non-tumor bearing brain images have been presented, the features of tumor and non-tumor MR files have been extracted through GLCM and watershed segmentation. The process is computationally less costly as compared to other techniques. Comprehensive literature review has been discussed along with the previously conducted research related to brain tumor segmentation and classification. Performance and accuracy of SVM, KNN and ANN have been evaluated which gave significant results. The approaches presented in this paper will aid significant application for addressing broader clinical challenges and will help

radiologist in diagnosing the tumor region with minimum manual operation and accuracy.

The performance of different algorithm presented in this paper are evaluated individually, a hybrid approach with ANN can produce better results. With hybrid approach more testing on larger dataset having different variety of MR images shall be performed. The proposed system can be extended to detection of tumors in other parts of the body such as lungs, bone cancer, kidneys. Since ANN produced better results so the proposed method can be used in conjugation with other methods like Wavelet, K-means, Fuzzy C means for enhancing accuracy and flexibility.

REFERENCES

- [1] M. Masroor Ahmed, Dzulkifli Bin Mohamad, "Segmentation of Brain MR Images for Tumor Extraction by Combining Kmeans Clustering and Perona-Malik Anisotropic Diffusion model.
- [2] Nagalkaar. V.J and Asole S.S, "Brain Tumor Detection using Digital Image Processing based on Soft Computing," Journal of Signal and Image Processing, Volume 3, Issue 3, Issn: 0976-8882, 2012.
- [3] Rajesh C.Patil, Dr. A. S.Bhalchandra, "Brain Tumor Extraction from MRI Images using MATLAB," International Journal of Electronics, Communication & Soft Computing Science and Engineering, Volume 2, Issue 1, ISSN: 2277-9477
- [4] Shivakumarswamy G.M, Akshay Patil.V, Chethan T.A, Prajwal B.H, Sagar.V.Hande (2016). Brain tumour detection using Image processing and sending tumour information over GSM.
- [5] Saha, B. N., Ray, N., Greiner, R., Murtha, A., & Zhang, H. (2012). Quick detection of brain tumours and edemas: A bounding box method using symmetry. Computerized medical imaging and graphics, 36(2), 95-107.
- [6] J. C. Dunn, A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters, Journal of Cybernetics, Vol. 3, No.3, 1973, pp. 32-57.
- [7] Dehariya, Vinod Kumar, Shailendra Kumar Shrivastava, and R. C. Jain. "Clustering of Image Data Set Using K-Means and Fuzzy KMeans Algorithms."Computational Intelligence and Communication Networks (CICN), 2010 International Conference on. IEEE, 2010.
- [8] P. Mohanaiah, P. Sathyanarayana, L. GuruKumar, "Image Texture Feature Extraction Using GLCM Approach", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.
- [9] Kshitij Bhagwat, Dhanshri More, Sayali Shinde, Akshay Daga, Assistant Prof. Rupali Tornekar, " Comparative Study Of Brain Tumor Detection Using K Means ,Fuzzy C Means And Hierarchical Clustering Algorithms " International Journal Of Scientific & Engineering Research , Volume 2,Issue 6,June 2013, Pp 626-632.
- [10] S.Roy And S.K.Bandoyopadhyay, "Detection And Qualification Of Brain Tumor From Mri Of Brain And Symmetric Analysis", International Journal Of Information And Communication Technology Research, Volume 2 No.6, June 2012, Pp584-588.
- [11] Tahir Sag Mehmet Cunkas, "Development Of Image Segmantation Techniques Using Swarm Intelligence",Iccit 2012.
- [12] Kalaiselvi, T., & Sriramakrishnan, P. N. P. (2016). ASimple IMAGE PROCESSING APPROACH TO ABNORMAL SLICES DETECTION FROM MRI TUMOR VOLUMES. Int J Multimed Appl, 8, 55-64.
- [13] V.Vani, M. Kalaiselvi Geetha (2016). Automatic Tumor Classification of Brain MR Images
- [14] Nikita V. Chavan, B.D. Jadhav, P.M. Patil (2015). Detection and classification of Brain Tumors
- [15] Kamal Kant Hiran, Rushi Doshi, (2013) An Artificial Neural Network Approach for Brain Tumor Detection using Digital Image Segmentation