Brain Tumor Detection and Classification from MRI using Machine Learning and Image Processing Novel Approach

[1]Melaku Eneayehu , [2]Rahul Joshi, [3]Foram Soni,[4]Betselot Yeulu,[5]Temesegan Walelign
[1]PG Scholar, [2]Asst.Professor, [345]PG Scholar
[1]Computer Science & Engineering,
[1]Parul Institute of Engineering & Technology, Parul University, Vadodara, India

Abstract : ***Brain tumor is currently a series problem for our world many people die because of abnormally growth of human red blood cell inside human skull. Detection and classification of brain tumor is too difficult because of the complex nature of brain. Now a day's different modern medical imaging technology is used to capture the clear image of our brain including soft tissue. Magnetic resonances Image (MRI), Posterior Emission Tomography (PET), X-Ray CT-Scan are some of the popular one. Among them MRI becomes the most common and popular because of free-radiation. Traditionally brain tumor segmentation and classification is done by radiologist that is vulnerable and prune to error, Medical Image Analysis is now becomes a hot research area in the field of Engineering and Medical, many research and method is proposed and currently also going on. In this paper a novel approach is proposed to detect and classify brain tumor using combination of image processing and machine learning algorithms. This proposed method is valid for all types of image type including the original DICOM (Digital Imaging of communication for Medicine). Image processing is mainly used for pre-processing step including skull removal, after skull removal the data is feed to unsupervised machine learning algorithm combination of K-means and fuzzy c-means segmentation, the next step is to apply Wavelet transform for feature extraction followed by Principal component analysis to reduce the dimensionality, next we evaluate around 13 features of statistical measurement, Kernel Support vector machine is used for classification of the tumor being benign and being malignant, the area of the tumor is also calculated, maximum accuracy is achieved from Gaussian radial basis with percentage of 92.6 %.***

IndexTerms - Tumor, Machine Learning, Kernel Support Vector Machine, PCA, DWT, DICOM, Benign, Malignant, Image Processing.

I. INTRODUCTION

Brain tumor or sometimes it can call neoplasm is when cells grow abnormally without control. In normal biological way new cell create to replace the work of other cell when that cell die, sometimes new cell created before die of other cell [1]. There are two types of neoplasm primary and secondary or benign and malignant, benign brain tumor is non-cancerous as well as non aggressive start to spread to brain the most well known cancer type that will spread to brain's are breast and lung cancer.

The classification as well as the grading system is given by WHO, according to their cell type and grade by visualize the cell [1]. MRI stands for magnetic resonance image it is non-invasive machine that works by generating radio wave without any radiation that affect our body for creating image [2]. MRI uses a powerful magnet where the power is measured in Tesla, the power of MRI machine that used for creating the image is reaches to 1.5T [3]. There are some terms that comes to in mind when we talk about MRI imaging technology, TR represents the repetition time, TE represents the Echo time, TR defines the successive repetition of Radio frequency energy, TE shows the delay time before Radio frequency energy, the most famous image sequence in MRI is T1-Wighted and T2-Weighted [4] . T1-weighted image is created by short TR time as well as short TE where as T2-wighted image in contrary created by long TR and long TE time [5].

In normal brain MRI T1-Wegihted sequence brightness is increasing in the order of CSF,GM and WM, but in contrary when we look T2-Weighted image the reverse is true brightness in increasing in the order of WM,GM and CSF, when we see the Tumorous image in T1-Weighted image it appears same or darker than others, but this is not always true example glioblastoma, in T2-Weighted image the tumorous part is high contrast or signal value than others, but this condition is not always true, in case of lymphoma tumor[6] this makes detection and classification of brain tumor difficult task. Brain tumor segmentation is very difficult task because, in different image modality brain tumor appears different in color and position[8]. This paper is organized in the following section in section 2 we will see related work that is done before in section 3 we see the proposed method which has its own sub-section section 4 will discussed about different performance evaluation in section 5 and 6 we draw the conclusion and future scope respectively.

II. RELATED WORK

In brain tumor detection until now many research is done and going on, in this section we will see the detail of some the methods that are proposed previously. Manisha, Radhakrishnan.B, Dr. L.P. Suresh 2017[7] proposed a method for extraction of tumor using edge detection, the weakness of this method is it only effective on high intensity image. Y. Megersa, Dr.G. Alemu 2015[6] propose for detection and segmentation of brain tumor using fast fuzzy c-means and hope-field neural network, The author mentioned that segmentation of brain tumor is very difficult task because, in different image modality brain tumor appears different in color and position. M.avula, N. Prasad,M.Prasad[8] proposed a method for bone tumor detection using mean pixel intensity, this method is good for T2-weighted image but, it got difficult for T1-Weighted image when the tumor is dark and same intensity with other tissue. R.Nouredine, K.Tarhini, S.Saleh[9] proposed a method for extraction and segmentation of brain injury from MRI using region growing algorithm, the weakness of this method is it needs a radiologist for selection of the seeded point. M.Rasoul[10] proposed a method for breast cancer detection using KNN, this proposed method use both image processing and machine learning. K. D.Mistry, B.J.Talati[11]proposed a method for bone tumor detection using combination of K-means and Fuzzy C-means clustering. S.Kare[12] propose a method using optimization technique like curve fitting and machine learning algorithm for detection of brain tumor. N.Afsha, S.Qurshi, S. Mujtiba [13] compares the three combination K-means, Fuzzy C-means and thresholding and finally they conclude combination of K-Means and Fuzzy C-Means gives a good result. R. Preetha and G. R. Suresh [14] proposed a method for detecting of brain tumor using throusholding and Fuzzy C-means clustering. Prof. S.Kumar[15] proposed a method segmentation of brain tumor using K-Means and DBSCAN. S.Gahann,N.Galkwar,N.Kulakarri, M.nekar[16] propose a method for brain tumor segmentation using fuzzy c-means clustering. A.Kavr[17] proposed a method for brain tumor detection using genetic algorithm. A.Singh and Parven in 2015[18]

proposed a method using fuzzy c-means and SVM. V.Surudhi, K.Sanjana, R.Saravanan, G.Santhosh, S.Kirubha in 2017[19] proposed a method for brain tumor detection using k-means algorithm. S.A.Kumar and T.C.Rao proposed a method using K-means clustering and morphological image processing. S. D. SABLE in 2015[20] proposed a method for breast-cancer detection using k-means and fuzzy c-means clustering.

III. PROPOSED WORK

The proposed method for brain tumor detection and classification using image processing machine learning is shown in figure 1 which consist of pre-processing, skull stripping, segmentation using combination of K-means and Fuzzy C-Means, feature extraction using Wavelet transform, followed by PCA for dimensionality reduction, after this step 13 statistical feature selected and feeds to kernel SVM for classification, finally the area of tumor will be calculated and the stage also declared.

A. Image Acquisition and Pre-Processing

Image acquisition, in this stage images are imported to our system, the image is in any image format, including original DICOM (Digital imaging and communication for Medicine) and other image format, our system support access to original medical image format. If the image is DICOM file format normally its grayscale image therefore it's not necessary to convert, but if the image is in other file format may be its RGB so we have to check the image format if the image is RGB it will convert to grayscale image. To do this we compare the image color channel, if it is more than one the image is not in gray scale format therefore conversion is needed.

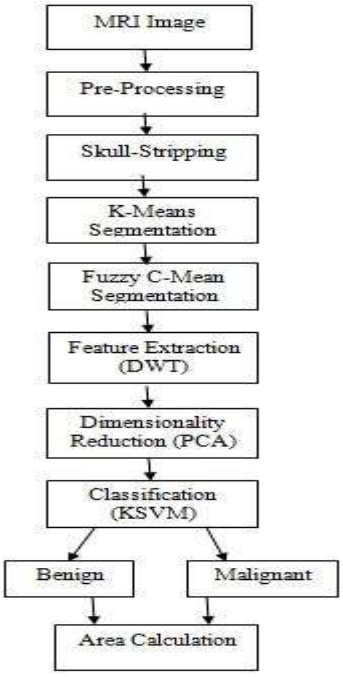


Figure1. Proposed Method

Always data is not getting easily as we want. Due to some case, in this case MRI image is noisy due to different reason, E.g. because of MRI machine or some transmission media. According to [21] MRI contains many noisy that will lead doctors to incorrect decision. Such as technician performance, MRI image by itself, the environment and patient clothing style. Therefore it is necessary to de-noise the image before using. According to[22]median filtering is more suitable for brain MRI image filtering than others like mean filter, Anisophrophic filter, Adaptive filter, Winner filter and KSL filtering. Additionally[21] noise in MRI may be introduced by, may be high magnetic field strength, pulsed radio frequency, i.e. what is the oscillation of pulse per second, radio frequency coil(receiver and transmitter), Bandwidth of receiver and voxel volume. Median filter is used to increase the visualization of image by removing noise from an image that will help for further decision or later pre-processing [23]. After Median filter the next step is Contrast enhancement before using the Brain MRI image we must increase the contrast of the image. Main limitation of MRI technology for brain diagnosis is it can't filter only the gray matter, white matter, cerebral fluid and tumor part rather it also adds the skull part and finally the images have low contrast [24].

B. Skull-Stripping

Brain is protected by strong cover called skull. During MRI imaging skull is one part of the image, but unfortunately we are not using the skull for brain tumor segmentation, therefore we have to remove the skull part the process is called skull stripping. Until now many research are going on separately i.e. only Target on how to remove the skull part from the brain or as an image pre-processing step. Many algorithms are developed but Morphological algorithms Erosion and dilation the most common and effective ways for skull removal but it is inefficient with time [25]. In this work we use mathematical Morphological operation.

C. Segmentation Technique

There are many segmentation techniques that allow extracting or detecting brain tumor. Clustering is a non-supervised machine learning method that allows to group objects based on some criteria sometimes clustering algorithm is called learning by observation rather than learning by example [26]. The most known clustering technique for image segmentation is K-means and Fuzzy C-means. Fuzzy C-Means clustering is sometimes known as by name called soft clustering that allows a single object can be grouped (clustered) into two or more cluster. Fuzzy c-means clustering nowadays used in many pattern and recognition filed [16 19 27]. Fuzzy is a way of manipulating data by providing partial membership to every pixel in an image. The value of membership is set range from 0 to 1[18]. It allows member of one cluster also member of other cluster.

K-Means clustering is one of the well known unsupervised machines learning clustering algorithm that allows clustering the query point with minimum distance measure from the query point.

D. Feature Extraction

In pattern recognition field feature extraction is a special type of dimensionality reduction that allows simplifying the most complex data. After image segmentation important feature will be extracted from segmented Image, many feature extraction algorithms are available and proposed for brain tumor detection in this methodology we apply Discrete Wavelet transform (DWT) followed by Principal Component analysis (PCA) for dimensionality reduction.

E. Discrete Wavelet Transform

DWT becomes the most popular and powerful method for medical image analysis starts from de-noising and for image compression. DWT is work by passing the signal in to a series of high and low pass filter, from low pass filter we get the approximation coefficient and from high pass filter we get the detail coefficient. Wavelet transform works generally by dividing the image signal in to different component [29], 2D-DWT becomes more popular in the field of Image processing. Fig2 shows single level decomposition of an image using DWT.

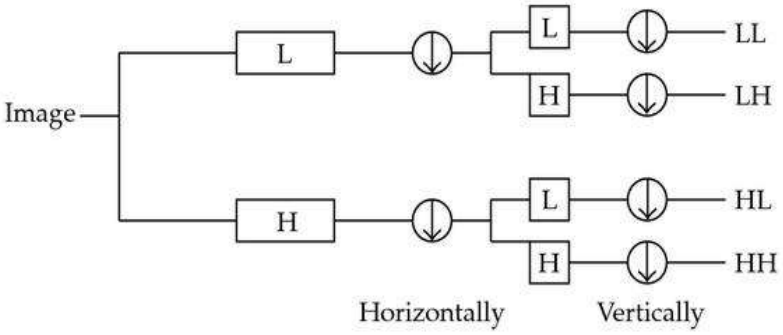


Figure 2. Single level DWT Decomposition for image

Where L-shows low-pass filter, H-represents High-pass filter, LL-approximation image that will decompose further if needed, LH, HL, HH represent the detailed coefficient for Vertical, Horizontal and Diagonal component respectively.

F. Principal Component Analysis

Many algorithms work better with low dimension and get trouble when they face high dimensionality data. When the number of feature that extract from the data in this case image, increases the time and space complexity is increasing therefore there must be some mechanism that handles dimensionality reduction, in this proposed method Principal component analysis will be implemented. PCA allows reducing the dimensionality of the image while keeping the information. It is one of the most powerful statistical methods. PCA has a lot of application in the field of digital imaging process, including classification of data, compression and facial recognition. PCA transforms X→Y, while X is matrix N*d to a matrix Y N*m. PCA concentrate to find the correlation among data, the correlation ranges from +1 to -1, where +1 indicates the positive correlation and -1 shows negative relation[30].

G. Support Vector Machine

SVM is a supervised machine learning model that used for regression and classification of linear as well as non-linear separable data. In non-linear separable it transforms to high dimensionality for finding a hyper plane (decision boundary) that will separate linearly. First invented SVM were linear SVM which used to separate linear data, later Kernel SVM was developed by M.Guyon and Vladmir N.vapnik for non-linear separable data it the technique we call kernel trick. In linear svm decision boundary calculate by a.x+b=0, a.x is a scalar product of a and x. [26].

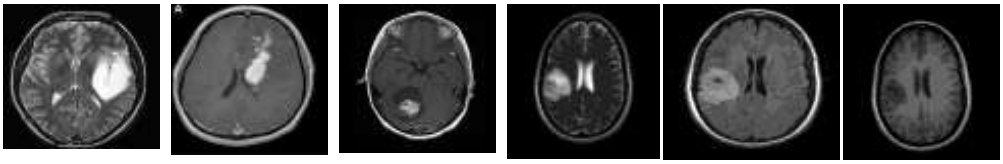
IV. EXPIRMENT RESULT

In the proposed method totally we use 205 images, from this 80% used for training and 20% used for testing; the first after different conversion and pre-processing is skull-stripping.

The next step is segmentation in this method K-means and fuzzy c-means segmentation implemented respectively the output of the segmentation result is shown below.

The next step is feature extraction, classification and area calculation of tumor for feature extraction we used 2D Discrete wavelet transform and the dimension is reduced using Principal component analysis 13 statistical is used for classification and three kernel function is used for performance evaluation.

A.



B.

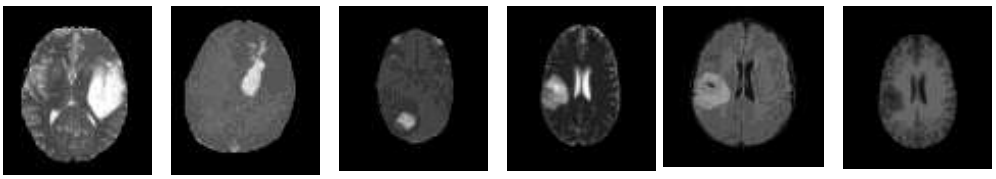


Figure 3. Shows the skull-stripping method that is done using mathematical morphological method in column A there is normal MRI brain image and in column B it shows the skull-stripped image its time complexity is high but it give good result, the first three images is benign tumor type and the second three images is malignant tumor type.

1.

2.

3.

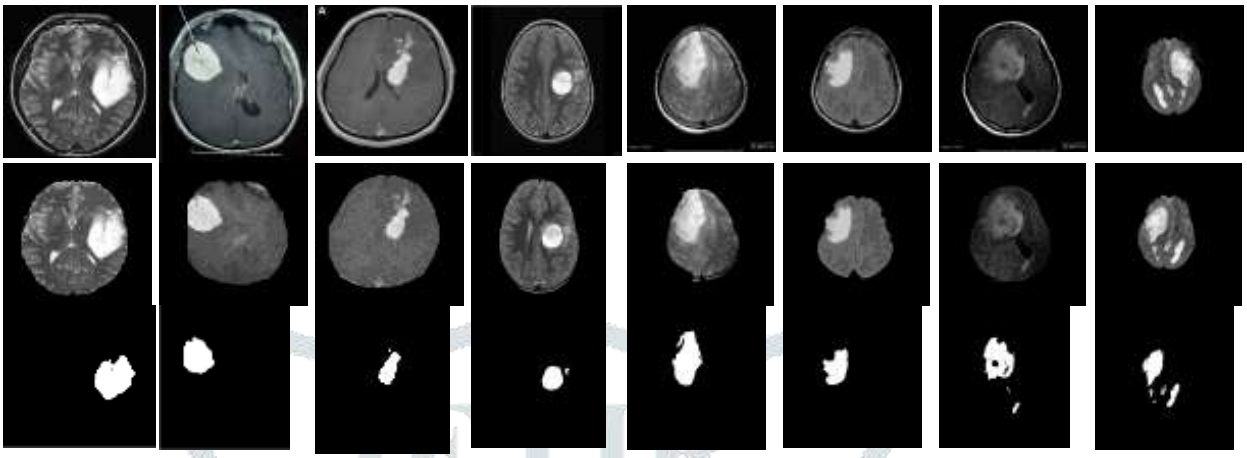


Figure 4. Shows the sample segmentation result in the proposed method in column a MRI T2-wighted scanned image shown but this method is also valid for T1-wighted image, in column b the skull-stripped image is shown, in the last column the segmented brain tumor image is shown.

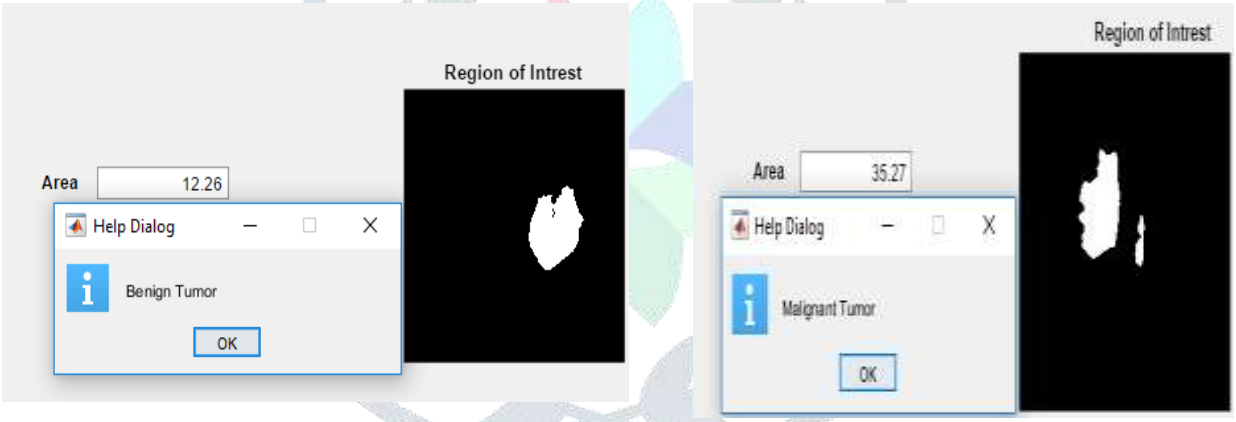


Figure 5. shows sample output for the classification and area calculation in column a it shows the tumor type is benign and the area is 12.26mm2 while in column b it shows the tumor type is malignant and the area is 35.27mm2, the area is calculated based on the formula (√P)*0.26458mm2, where P represents the number of white pixel in the binary image and 0.26458 because 1pixel=0.26458mm.

V. PERFORMANCE EVALUATION

Since when we build our model, during classification the support vector machine classifier trained by training dataset, sometimes what happen it gives better result only for trained data, when we test unseen data it will fail this condition is called over fitting. Over fitting is the main problem while developing supervised model, to avoid this we used K-fold cross validation, which divide the original sample into equal K-sized subsamples, then K-1 subsamples is used for training data, and only 1k (subsamples) is used for validating(testing). In this proposed method we used 10 K-fold cross validation, which use 9 for training and 1 for validation, the process will iterate till k-times. In this paper totally we used 205 images, 160 images used for training and 45 images used for testing purpose from 45 image the system correctly classify 37 as benign and malignant while 8 image classified incorrectly while we use 3 kernel function namely Gaussian Radial Basis, linear kernel, polynomial kernel (Homogenous and Inhomogeneous) from all this Gaussian Radial Basis gives a high accuracy from all with percentage of 92.6% Accuracy.

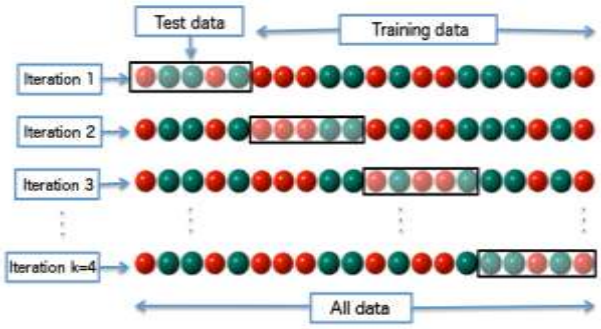


Figure 6. K-fold cross validation with K=4

VI. CONCLUSION AND FUTURE WORK

Brain tumor or neoplasm is a very serious problem in the past, now and future. Medical imaging allows to see different parts of our body clearly specially MRI. Medical image analysis currently is hot topic for research because of availability of huge data from time to time. Many methods are developed previously, in this proposed method we show how to segment and classify brain tumor from MRI using combination of k-means and fuzzy c-means segmentation and wavelet transform for feature extraction, principal component analysis is used for dimensionality reduction we get 87.8% accuracy, in this work we try to include both T1 and T2-wighted image. Many works can be extend from this work starts from improving skull removal step, time complexity since we used fuzzy c-means it takes high computational time, so reduce computational time in one work that can be extend, finally 87.8% accuracy is not enough, increasing accuracy and analysis 3-D can be extend from this work.

REFERRNCE

[1] M. Eneayehu and R. Joshi, "A survey on various brain tumor detection method using Image Processing and Machine Learning," Parul University International Conference on Engineering & Technology: Smart Engineering february 16-17, 2018, Parul University ,Vadadora, Gujarat, india.

[2] M. Eneayehu and R. Joshi, "Brain tumor detection and segmentation from MRI using unsupervised machine learning algorithm," International Conference on Electrical, Electronics, Computers, Communication, Mechanical and Computing (EECCMC) 2018 28th & 29th January 2018,Priyadarshini Engineering College, Chettiyappanur, Vaniyambadi - 635751, Vellore District, Tamil Nadu, India.

[3] https://en.wikipedia.org/wiki/Magnetic_resonance_imaging Accessed on September 13 2017.

[4] Yehualashet Megersa, MSC Thesis, "brain tumor detection and segmentation using hybrid intelligent algorithms: Design and Implementation", Addis Abeba University, Ethiopia, November 2012.

[5] http://casemed.case.edu/clerkships/neurology/Web%20Neurorad/MRI%20Basics.htm Acccessed on September 14 2017

[6] Y. Megersa and G. Alemu, "Brain tumor detection and segmentation using hybrid intelligent algorithms," AFRICON 2015, Addis Ababa, 2015, pp. 1-8.

[7] Manisha, B. Radhakrishnan and L. P. Suresh, "Tumor region extraction using edge detection method in brain MRI images," 2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT), Kollam, 2017, pp. 1-5.

[8] Avula, Madhuri, Narasimha Prasad Lakkakula, and Murali Prasad Raja. "Bone Cancer Detection from MRI Scan Imagery Using Mean Pixel Intensity." IEEE2014 8th Asia Modelling Symposium 978-1-4799-6487-1/14, 2014. 141-146

[9] Rasha Nouredine, khalil Tarhini, Soha saleh"Segmentation and Extraction of brain injury lesion from MRI image Matlab Implementation" IEEE 2015 International Conference on Advances in Biomedical Engineering (ICABME) ,45-48

[10] Moh'd Rasoul "Breast cancer detection using K-nearest Neighbhr Machine learning Algorithm" IEEE 2016 9th International Conference on Developments in eSystems Engineering, 2016, 35-39

[11] Kaupali D.mistry, Bisal J.Talati "Integrated Approach for Bone Tumor Detection from MRI scan Imagery " IEEE International Conference on Signal and Information Processing (IConSIP),2016,1-5

[12] Sneha kare"Optimization technique Curve fitting and Machine learning used to detect Brain tumor in MRI" IEEE IEEE International Conference on Computer Communication and Systems(ICCCS 114), Feb 20-21, 2014, Chennai, INDIA ,254-259

[13] Nailah Afsha, Sharma Qurshi, Seyed mujtiba "Comparative study of Tumor Detection Algorithm" International Conference on Medical Imaging, m-Health and Emerging Communication Systems, 2014, 251-256

[14] R. Preetha and G. R. Suresh, "Performance Analysis of Fuzzy C Means Algorithm in Automated Detection of Brain Tumor," 2014 World Congress on Computing and Communication Technologies, Trichirappalli, 2014, pp. 30-33.

[15] Prof. Samir Kumar "segmentation of brain tumor from MRI image- Analysis of K-means and DBSCAN clustering " IJRES Volume 1 Issue 1 ǁ May. 2013 ǁ PP.48-57

[16] Swati etal "Detection of Possibility of brain tumor using image segmentation " Vol. 3, Issue 4, April 2015,2701-2704

[17] Amnapreet Kavr "Overview of tumor detection using Genetic Algorithm" , IJIET, Vol 2 Issue 2 April 2013,348-352

[18] Amit singh, Parveen"Detection of brain tumor in MRI Images Using Combination of fuzzy C-means and SVM" IEEE 2nd International Conference on Signal Processing and Integrated Networks (SPIN),98-102

[19] V.Sagar Anil Kumar, T.Chandra Sekhar Rao "Brain tumor extraction by K-means clustering Based on morphological image processing" International Journal of Engineering and computer science (IJECS),Vol3, Issue 10, October 2014 page no. 8539-8546

[20] SNEHALI D. SABLE, "implementation and analysis of K-Means and Fuzzy C-means clustering techniques" Proceedings of 26th IRF International Conference, 10th May 2015, Pune, India, ISBN: 978-93-85465-09-3,66-69.

[21] S.Valishali,Dr.K.Kishan Rao,Dr.G.V Subba"A Review on Noise Reduction Methods for Brain MRI Images" International Conference on Signal Processing and Communication Engineering Systems ,2015,363-365

[22] Anisha, S.R PG Scholar, Dr J Venugopala Krishnan "Comparison of Various Filters for Noise Removal in MRI Brain Image"International Conference on Futuristic Trends in Computing and Communication (ICFTCC-2015),68-73

[23] Ruchi D. Deshmukh, "Detection of brain tumor with Filtering Tecniques" International Journal of Computer Science Engineering and Technology (IJCSET) | April 2014 | Vol 4, Issue 4,133-136
[24] L. Kapoor and S. Thakur, "A survey on brain tumor detection using image processing techniques," 2017IEEE 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, Noida, 2017, pp. 582-585.
[25] Sajjad Mohsin, Sadaf Sajjad, Zeeshan Malik, and Abdul Hanan Abdullah "efficient way of skull stripping in MRI to detect brain tumor by applying morphological operator, after detection of false background" International Journal of Information and Education Technology, Vol. 2, No. 4, August 2012.
[26] Jiawei Han and Micheline Kamber ;data mining concept and techniques; 2nd edition Morgan Kaufmann Publishers, 2006.
[27] Nailah Afshan, Shaima Qureshi, Syed Mujtiba Hussain "study of different brain tumor MRI image segmentation tecnique" IEEE International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom) 2014,251-256
[28] Zijun Zhang Presention on K-Means Algorithm, Cluster Ananlsysis in Data-Mining. 2012
[29] D.Gupta and S.Choubey "Discrete Wavelet Transform for Image Processing" International Journal of Emerging Technology and Advanced Engineering, Volume 4, Issue 3, March 2015
[30] https://www.slideshare.net/foofiM/principal-component-analysis-45838432 accessed on January 2018.