Hybrid K-Means and Watershed, Texture Machine Learning Approach to Predict Brain Tumor

A.Benazir1, E.Ramaraj2
1(computer science, Alagappa University, and karaikudi
Email: benaabdul07@gmail.com)
2 (computer science, Alagappa University, and karaikudi
Email: dcsalagappauniversity@gmail.com)

Abstract:

Image processing is utilized in the analytic machine for tumor detection. Brain tumor is a major issue in the restorative field. Image segmentation is used to detect tumor from the brain.MRI scanned images are utilized to identify brain tumor with the help of existing k-means segmentation. Various distinct schemes such as pre-processing, MRF&CRF used to enhance the quality of edge detection and the combination of k-means and watershed, texture segmentation. The proposed method described in this paper is used to improve the image clarity. A machine learning concept and proposed KWT (K-means, Watershed, texture) method are used to predict the brain tumor with better accuracy, sensitivity and specificity. In this paper, comparative analysis is made for existing and proposed k-means methods. The quality metrics such as PSNR&MSE yields better accuracy than the exiting methods.

Keywords — **Machine learning, Canny Operator, MRF and CRF, k-means, watershed and texture segmentation.**

I. INTRODUCTION

Image processing is the capacity to change over an image into a digital format, and it is used to accomplish some functions in the image, therefore exact image that can be archived through it. Data can be better comprehended through pictures. Generally machine learning target on removing noisiness from an image after that significant data is utilized to manage different tasks. Image segmentation technique enables digital image processing for multidisciplinary medicinal image segmentation in automated or semi-automatic detection of limits inside 2D or 3D images. Recent research concentrates on the pattern mining methodology to predict defect portion from the digital image input [16].Decision about the defect portion in computation process is based on the positive and negative parameters of the resultant values [17]. Tumor is one of the major health issues in human life. Infected cell prediction helps to find the tumor in early stages [18]. The brain tumor is an unexpected growth of cells within the brain that prevents brain activities such as muscle control, sense, memory, and other normal physical functions. There are two principal sorts of brain tumours, primary and metastatic. The primary tumor starts and tends to stay in the brain, whereas metastatic brain tumor begins as cancer elsewhere in the body and spreads to the brain. There are two kinds of primary tumours Beginning and Malignant (mal=" bad", lgnis=" fire"). There are no cancer cells in a beginning tumor, while malignant tumours contain cancer cells. MRI (Magnetic Resonance Imaging) Scan is a radiographic technique that uses radio waves, magnetism and the internal structure of a brain tumor using a computer.

II. RELATED WORK

J.Vijay, J.Subhashini, this paper portrays a productive strategy for mechanized brain tumor segmentation to extract tumor tissues from MR images [1]. Anupurba Nandi, 2015, This paper uses K-Means clustering, where some of the detected tumor irregularity is then corrected by using image drivers with basic image processing techniques to achieve the goal of differentiating tumor cells from normal cells[2]. Digvijay Reddy, Dheeraj, Kiran,

Bhavana.V and Krishnappa H.K, April 3-5, 2018, in this work, Dycom [3] has been assigned to magnetic resonance imaging input and attempts to extract tumor cells from the input image. The pre-processing technique is utilized to discard noise from the image. For this image, K-clustering is used and from this cluster image, the skull is removed using morphological functions to easily identify tumor cells. Miss. Shrutika Santosh, Prof. Akshata Raut Hunnur, Prof. Swati Kulkarni, 2017, this study describes the diagnosis of a brain tumor by thresholding. The proposed technique can be effectively used to recognize and remove MRI images of brain tumors obtained from a patient's database[4].ManuGupta,Prof. B.V.V.S.N.Prabhakar Rao, Dr.Venkateswaran Rajagopalan, 2016, This paper depicts a new framework for brain tumor diagnosis and its classification based on high-order statistical system features such as kurtosis and curvature with selected morphological features. To distinguish high-grade (Hg) tumors from low-grade (LG) tumors, these appearances were exact from separated tumor T2-weighted brain MR images [5]. Swap nil R. Telrandhe1, Amit Pimpalkar2, Ankita Kendhe3, This study, proposes adaptive brain tumor detection, image processing is used in clinical instruments for tumor detection cannot be identified tumor regions only with MRI images, so K-Means section is used [6].

III. PROPOSED FRAMEWORK

The major reason for this paper is to realize its region point analysis of malignant growth and tumor is used to treat the disease. It has been concluded from the research of most of the developed countries that 73% of brain tumor deaths occur when compare to other cancers ie.47%. It is predicted that 23,380 men (12,820 men and 10,560 women) in the U.S. will be affected with brain cancer. It is predicted that 14,320 male (8,090 men and 6,230 women) will die of this brain tumor in 2014[7]. Around 22,850 threatening tumours of the brain (12,900 in man and 9,950 in women) are analysed. Around 15,320 individuals (8,940 man and 6,380 women) die due to brain and spinal tumours. Overall, the likelihood of a person developing a malignant tumor in their lifetime is less than 1% in 2015 (1 in 140 for a man and 1 in 180 for women). This is a terribly difficult disease to diagnose. In the present work, MATLAB tool is implemented. It can be used to improve brain tumors detection more accurately. The image segmentation bottle is done by various techniques. First MRI scan Images are taken and the process can be done by the following steps. The first step is image acquisition. Here the MRI brain images are converted to the grayscale images. In the pre-processing phase, the input MRI images are tested to identify whether the MRI image is normal (benign) or abnormal (malignant). Then the Threshold Method is used to separate the background and ROI (Region of Interest) ie.tumor portions. The tumor obtained doesn't have exact edges, so the detection method of the canny edge is used. The noise is then eliminated by image-based de-noise. In the clustering Method combination of MRF and CRF was used. Finally, the main objective of this work is a combination of K -means and Watershed segmentation, which can be used to overcome the over-segmentation and texture segmentation to identify the tumor region of the brain MRI image as input. The proposed system is shown below in the block diagram.

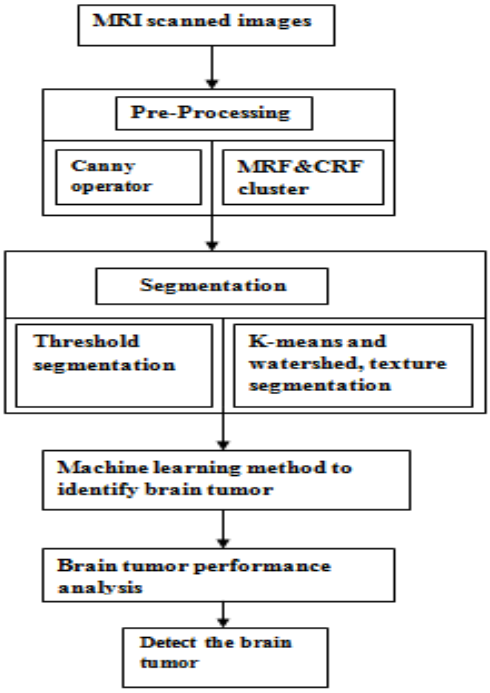


Fig 1: Flow of Proposed Method

A. Image Acquisition

Brain tumor Images that are obtained by using the MRI scan images are converted from RGB to Grey level. In Weighted method gives a grey colour image by using the formula (0.3*R)+(0.59*G)+(0.11*B) i.e. In terms of intensity, 30% red, 59% green and 11% blue give a clear gray scale image.

B. Pre-Processing

In this pre-processing phase, the input MR images are tested to identify whether the beginning or malignant tumor is present in the input image. Gray scale images are shown in a two-dimensional matrix from 0 to 255, where 0 display the total black and 255 white, so that the background and ROI can be separated, and later edges are detected using canny operators. Noise is eliminated by deformation-based denoising-non-local methods [8] [9], which are efficient and give better conclusion compared to other filters. Next, in the clustering methods, the combination of MRF and CRF reduces both the problem of overlapping and noise with high accuracy [10] [11].

C. Processing Stage

The analysis of brain tumor involves two major steps:

a. Image Classification
b. Image Segmentation

In image classification, brain tumor detection is based on similarity measurement. In the Image Segmentation diagnosis of Brain, the tumor is based on size and shape.

D. Post-processing Stage

Segmentation processing is done by the following methods. In the computer view, segmentation image is a way of processing the computerized image into different parts (a set of pixels, it is called super pixels). It is used to find objects and borders in images.

IV. METHODOLOGY

A. Threshold Segmentation

The Optimal Threshold is a simple but effective tool to separate the ROI and background.ie this segmentation is achieved by organizing all the pixels into one class and all another pixel into another class. The threshold value will change over the gray scale image into a binary image format which is utilized to select the threshold value [12] [13].

B. K-means and watershed segmentation

The Watershed segmentation is a great way to compile pixels of an image dependent on their intensity. It is a mathematically morphological operating tool and it is used for checking the output rather than input segmentation techniques because of its drawbacks over-segmentation and under segmentation. A combination of K -means and watershed segmentation overcomes the segmentation technique [14].

- ***Tumor identification***

Here, essentially filtered images of the brain are used. As scanned digital image of the brain, is taken that particular image ‘n’ and calculate the height and width and number of pixels of that image ‘n’ is converted to gray scale. Hence after this it is converted into 3d i.e. pixels of that image are converted to 3d ‘n’ clustering concept is applied to that image. During clustering, it is difficult to select the cluster size because K-means computation requires a finite number of clusters. The first mean value for the clustering has to be defined. After the clustering process is done, in reconstruction, finally the output is obtained. Where it will be clearing known whether that image contains a tumor or not. The accuracy of finding a tumor has been increased by using an automated system.

Watershed segmentation is a gradient-based segmentation technique. It considers the slope map of the picture as an intervention map. It segments the image as a dam. The segmented regions are called catchment basins. The Water segmentation solves a variety of image segmentation problems. This is reasonable for high-intensity values. Watershed segmentation is caused by over-segmentation. To control the segmentation, the marker-controlled watershed segmentation is utilized. The Sobel operator is appropriate for edge detection.

The Sobel masks in the matrix form are:

$$Mx=\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \quad My=\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

The equation of inclination magnitude utilized in the marker constrained watershed segmentation is

$$M=\sqrt{Mx^2+My^2}$$

$$\text{Angle, } \Theta=\tan\text{-}1\frac{mx}{my} \quad \ldots\ldots\ldots\ldots\ldots (1)$$

c. Texture Segmentation

The texture is a normal repeat of an element or shape on a surface that has some amount of variability in the origin and relative position of the element. The proposed system utilized two strategies for extracting the texture highlights. The principal technique depends on the main request histogram, which is a neighbourhood in nature and the subsequent strategy depends on the co-occurrence, known as the second-request framework highlight. Demonstrates the subtleties of the element extraction process. 1. First-request histogram-based highlights 2.Co-occurrence matrix based features: Histogram based features are local in nature. These highlights don't think about spatial data into thought.

So, for this reason, gray level spatial concurrence matrix HD (i, j) based highlights are characterized which are known as second-request histogram based highlights.

$$\sum_{I=0}^{G-1}\sum_{J=0}^{G-1}[p(i,j)]^2 \qquad (2)$$

$$\text{correlation: }\sum_{I=0}^{G-1}\sum_{J=0}^{G-1}\frac{p(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y} \qquad (3)$$

$$\text{Intertia: }\sum_{I=0}^{G-1}\sum_{J=0}^{G-1}(i-j)^2 2p(i,j) \qquad (4)$$

$$\text{Absolute : }\sum_{I=0}^{G-1}\sum_{J=0}^{G-1}|i-j|\ p(i,j) \qquad (5)$$

$$\text{Inverse Difference: }\sum_{I=0}^{G-1}\sum_{J=0}^{G-1}\frac{p(i,j)}{1+(i-j)^2} \qquad (6)$$

$$\text{Entropy: } H=-\sum_{I=0}^{G-1}\sum_{J=0}^{G-1}p(i,j)\ \log_2[p(i)] \qquad (7)$$

$$\text{Maximum Probability: }\max_{i,j} p\,(i,j) \qquad (8)$$

Calculation of the tumor area. The field of the tumor region is determined by the accompanying condition: Tumor area= A total group of pixels in the tumor area (5) A= V x H (6). Where A=the region of every pixel H=horizontal measurement of the image V=vertical measurement of the image H=1/horizontal goals of the image V=1/vertical goals of the image.

V. MRI IMAGE DATABASE AND READING

MRI images are extricated for brain tumor examination from the BRATS database accessible on the web. BRATS the database is a gathering of images identified with different diseases for research and examination purposes. Accessible MRI the data is as nii.gtz, .hdr and .img. The image data ought to be read as the essential advance of the activity in MATLAB. SPM, MRI, and Micron are the devices to read this and File Viewer.MRI data is very noisy due to high magnetism powerless to head development and others natural conditions.

Consequently, spatial filters MRI data ought to be balanced with the end goal of investigation. MRI got from MRI data is a lot of image successions (outlines) played over some undefined time frame. X-ray data is loaded into the MATLAB structure utilizing mad command. The accompanying images show MRI images, read them utilizing the MATLAB software. The use of benchmarking methods and methodology for tumor segmentation has reached extensive accuracy.

VI. RESULTS AND DISCUSSION

The segmentation is the path toward isolating an image into districts with comparable properties, for example, gray level, colour, texture, brightness, and differentiation. The major plan is to detect the tumor present in the patient brain, which is very useful for the diagnosis, which is very easy to take treatment for the survival of the patients. For the experiment, 20 brain tumor images obtained from the clinical laboratory are included to test the code. Describe the whole process of brain tumor detection. The result for the K-Means, Watershed and Texture segmentation are shown below and table.1 also clearly shows that it is implemented on different images taken from the laboratory for this research work. The BRATS image information and manual explanations continue being straightforwardly available through an online appraisal framework as an advancing benchmarking asset. The accuracy, sensitivity and specificity for the proposed strategy bottle be registered to utilize the conditions 9, 10, 11 individually accuracy is the probability that an

analytic test is right to the Overall Accuracy, Sensitivity and Specificity measures are used to ensure the quality in classification process [15].Accuracy, sensitivity, and specificity of the proposed ones calculate the method using equations 9, 10 and 11 respectively –

The probability accuracy that the diagnostic check is proper

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

The sensitivity (true positive fraction) is the probability that the diagnostic check is positive because the person has the tumor disease.

$$Sensitivity = \frac{TP}{TP+FN} \quad (10)$$

The specificity (true negative division) is the likelihood that a symptomatic test is negative, given that the individual does not have the disease.

$$Specificity = \frac{TN}{TN+FP} \quad (11)$$

TABLE I

LIST OF PERFORMANCE MEASURES

Statistical Test	Accuracy (%)	Sensitivity (%)	Specificity (%)
K Means Method	**86.6%**	**90.9%**	**75%**
Proposed Hybrid Method	**99.80%**	**94.80%**	**89.98%**

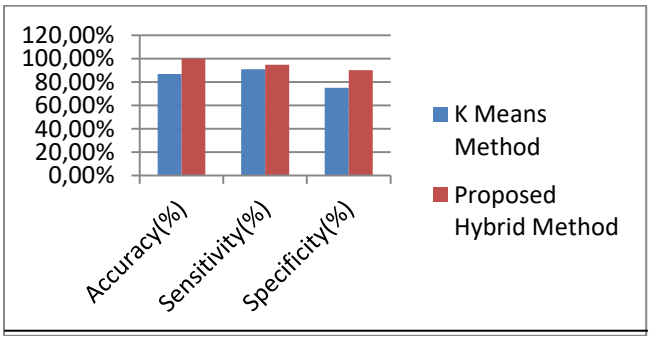
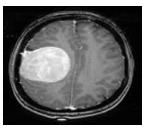
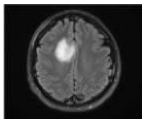


Fig 2: Analysis of proposed hybrid method

There are two common error options, namely the peak signal-noise ratio (PSNR) and the mean square error (MSE) shown in table1 above. calculation of MSR and PSNR have demonstrates the accuracy of the proposed algorithm and Combining a number of good methods using a hierarchical majority of votes provided segments that relied on each individual methodology, indicating remaining opportunities for further formal improvements .

TABLE II

	IMAGES	PSNR	MSE
Anisotropic Filter, Threshold Segmentation		**21.6398**	**64.6555**
Hybrid KWT method		**3.5273**	**85.3807**

PERFORMANCE ANALYSIS

VI. CONCLUSIONS

Present work proposes an effective technique for mechanized brain tumor segmentation on datasets standard benchmark. In this paper, an automated system has been made for ordering start and abnormal tumors in brain MRI images using the classifier. The accuracy of brain tumor analysis in MRI patients is very promising; it is utilized to classify tumor types according to the clinical diagnosis system. Along these lines, the objective of applying k-means and watershed combination segment are used to defeat the segmentation of surface and morphological administrators. The output image demonstrates the tumor cells which have been isolated from the solid cells. The achievement rate of the proposed algorithm represents 99.80% (true positive and true negative), 94.80 %(true positive), 89.98 %(true negative) results.

ACKNOWLEDGMENT

This article has been written with the financial support of RUSA-Phase 2.0 grant sanctioned vide Letter No.F.24-51/2014-U, Policy (TNMulti-Gen), Dep.of Edn.Govt. Of India, Dt.09.10.2018.

REFERENCES

1. *Vijay, J., and J. Subhashini. "An efficient brain tumor detection methodology using K-means clustering algoriftnn." 2013 International Conference on Communication and Signal Processing. IEEE, 2013.*

2. *Nandi, Anupurba. "Detection of human brain tumour using MRI image segmentation and morphological operators." 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS). IEEE, 2015.*

3. *Reddy, Digvijay, V. Bhavana, and H. K. Krishnappa. "Brain Tumor Detection Using Image Segmentation Techniques." 2018 International Conference on Communication and Signal Processing (ICCSP). IEEE, 2018.*

4. *Hunnur, Miss Shrutika Santosh, Akshata Raut, and Swati Kulkarni. "Implementation of image processing for detection of brain tumors." 2017 International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2017.*

5. *Gupta, Manu, BVVSN Prabhakar Rao, and Venkateswaran Rajagopalan. "Brain tumor detection in conventional MR images based on statistical texture and morphological features." 2016 International Conference on Information Technology (ICIT). IEEE, 2016.*

6. *Telrandhe, Swapnil R., Amit Pimpalkar, and Ankita Kendhe. "Detection of brain tumor from MRI images by using segmentation & SVM." 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave). IEEE, 2016.*

7. *Shi, Jichuan, and Hong Zhang. "Adaptive local threshold with shape information and its application to object segmentation." 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO). IEEE, 2009.*

8. *Ekin, Ahmet. "Pathology-robustmr intensity normalizationwith global and local constraints." 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, 2011.*

9. *Zhang, Xiaoyan, et al. "An image segmentation method based on improved watershed algorithm." 2010 International Conference on Computational and Information Sciences. IEEE, 2010.*

10. *Tran, T. N., Wehrens, R., & Buydens, L. M. (2005). Clustering multispectral images: a tutorial. Chemometrics and Intelligent Laboratory Systems, 77(1-2), 3-17.*

11. *Chen, W., Ding, M., Miao, Y., & Luo, L. (2011, December). Ultrasound image denoising with multi-shape patches aggregation based non-local means. In 2011 International Conference on Intelligent Computation and Bio-Medical Instrumentation (pp. 35-38). IEEE.*

12. *Hossam, M. Moftah, Aboul Ella Hassanien, and Mohamoud Shoman. "3D brain tumor segmentation scheme using K-mean clustering and connected component labeling algorithms." 2010 10th International Conference on Intelligent Systems Design and Applications. IEEE, 2010.*

13. *Li, Gang, and Youchuan Wan. "Improved watershed segmentation with optimal scale based on ordered dither halftone and mutual information." 2010 3rd International Conference on Computer Science and Information Technology. Vol. 9. IEEE, 2010.*

14. *Perona, Pietro, and Jitendra Malik. "Scale-space and edge detection using anisotropic diffusion." IEEE Transactions on pattern analysis and machine intelligence 12.7 (1990): 629-639.*

15. *Kr, Sivabalan & E, Ramaraj. (2017). Band Value Based Reflective Image Classification Method to Classify the Satellite Image Environment. International Journal of Engineering and Technology. 9. 3630-3635. 10.21817/ijet/2017/v9i5/170905323.*

16. *E, Ramaraj & venkatesan, n. (2019). an efficient pattern mining analysis in health care database.*

17. *E, Ramaraj & Venkatesan, N. (2008). Positive and Negative Association Rule Analysis in Health Care Database.*

18. *R. Mathumitha, Dr. A. Padmapriya(2016) "Prediction of Disease by Analysing the Cell Image of Human" International Journal of Advanced Research in Computer Science and Software Engineering. Volume 6, Issue 6, PP 766 -769.*