RESEARCH ARTICLE | SEPTEMBER 05 2023

Early identification of brain tumors from MRI images using fuzzy C means clustering

Karthik Veeravalli; Lokesh Kota; Ankush Kappe; Rahul Kamishetty; R. Soujanya

AIP Conference Proceedings 2754, 070003 (2023)
https://doi.org/10.1063/5.0161712

Early Identification of Brain Tumors from MRI images using Fuzzy C Means Clustering

Karthik Veeravalli[1,a)], Lokesh Kota[1,b)], Ankush Kappe[1,c)], Rahul Kamishetty[1,d)], R.Soujanya[1,e)]

[1]*Gokaraju Rangaraju Institute of Enigneering & Technology, Hyderabad, INDIA*

[a)]*Corresponding author:karthikveervalli@gmail.com*
[b)]*lokeshkota2101@gmail.com*
[c)]*ankush44.k@gmail.com*
[d)]*Rahulkamishetty01@gmail.com*
[e)]*soujanya96@gmail.com*

Abstract. Image processing plays a significant role in applications like computer vision, surveillance and security, video processing, pattern recognition, and more recently, it has gained massive backing in the healthcare industry. A tumor is an abnormal cell growth or mass in parts of the body. Early detection of tumors can have a significant impact on the outcome. Some tumors are benign (non-cancerous), while others are cancerous (malignant). Conventionally MRIs are preferred for diagnosing tumors because they produce more detailed images; however, CT scans and PET scans are also used for the purpose. However, the task of identifying, segmenting, and detecting the infected area is time-consuming and tedious. This research focuses on brain tumor detection and proposes a method based on noise removal, followed by image enhancement techniques. Subsequently utilizing image segmentation procedures and edge mapping methods to identify and chart the delicate edges of an MRI or CT scan image could drastically reduce the amount of time medical professionals need to study and comprehend tumors.

(Keywords- tumor, image processing, canny edge detection, BCET, healthcare.)

I. INTRODUCTION

The detection of brain tumors is one of the most critical tasks in medical image analysis. The term "brain tumor" describes a condition in which the number of cells in the human brain grows abnormally. It develops from nerves that emerge from the brain and brain cells and blood vessels in the majority of cases[1]. The pressure inside the skull can arise when tumors, whether cancerous or non-cancerous, grow more prominent. This can cause brain damage, which is potentially fatal[2]. The majority of medical establishments now use a categorizing system developed by the World Health Organization (WHO) to identify brain tumors. The World Health Organization divides brain tumors into four types based on cell origin and behavior, from the least aggressive (benign) to the most aggressive (malignant)[3].

Image processing is one of today's most progressively evolving technologies. The application of operations to an image to enhance it or to mine/extract valuable data and knowledge is known as image processing. It is a type of signal processing in which an image is used as input, and either the image itself or its characteristics/features are output[4].

There are currently two types of segmentation-based tumor identification and estimation methods available: region and contour based methods[5]. The region-based plans look for groups of pixels that are similar in some way and determine interest regions. Commonly, these methods rely solely on local information for every pixel and ignore shape and boundary-related information. The Fuzzy C Means (FCM) clustering technique is typically used for segmentation tasks; however, this paper proposes a hybrid approach that combines FCM clustering with the Canny edge detection technique for identifying tumor margins for various cases of brain tumors with the formation of contour representation with color mapping of brain tumor.

II. LITERATURE REVIEW

Accurate identification and diagnosis of human illness are critical in today's globalized world. More advanced technologies are required to aid this progress as medical science progresses.

Image processing techniques have been developed to assist doctors in detecting and identifying brain tumors. Some early efforts yielded promising results, but they also had some drawbacks [4].

One of the pioneering works in recent times, researched by Saurabh Kumar et al., the authors proposed two different methodologies to segment a tumor from the MRI scans. The model used one clustering technique and one segmentation technique, namely, Self-Organised Maps (SOM) Clustering and SVM classification [6]. However, the use of SOM makes the model computationally expensive. It also heavily relies on the weights assigned to each neutron, and when presented with either too little or excessive data, it struggles to generate accurate results[7]. Another such study conducted by Jinal A. Shah et al. proposed a model that uses FCM and K-Means for the segmentation process to detect brain tumors using MRI scans as input [8]. However, during the preprocessing phase, the image obtained is comparatively less clear due to loss of clarity. This problem was overcome in the present research using BCET and AMF, which enhance the image after preprocessing and retain the critical information from the MRI scan. The process of contrast enhancement/image enhancement is required to draw attention to the areas of interest during medical image processing. The Balance Contrast Enhancement Technique (BCET) increases contrast and highlights the area of interest. The image's contrast can be stretched or flattened without affecting an input image's histogram[9].

BCET is gleaned from a parabolic function that is calculated from the given input. Its general form is given as,

$$\boldsymbol{y = a(x - b)^2 + c}$$

Where the three coefficients 'a,' 'b,' and 'c' represent the minimum, maximum, and mean values of the output image, respectively.

In this [10] study conducted by Aaswad Sawant et al., the researchers focused on using TensorFlow to detect brain cancer using MRI in their research. The model implemented convolutional neural network architecture with five layers and a softmax activation function in the final dense layer. In the pooling layers, the researchers used the max pooling feature. Although the results were satisfying, certain drawbacks can be associated with the work due to the utilization of max pooling. In this technique, the pooling operator considers only the maximum (most valuable) constituent from the pooling area and ignores all other factors. If the majority of the elements in the pooling area are of high magnitude, the distinguishing components disappear once the maximum pooling operation is completed. In one other paper by Mohammed Elmogy et al. [11], the

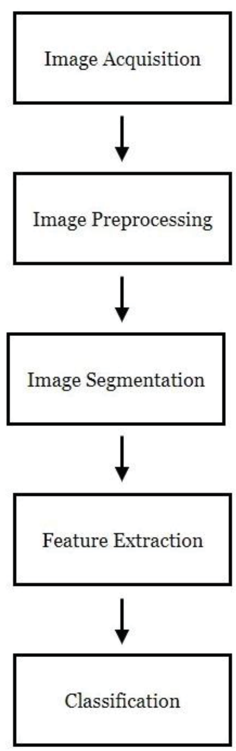

FIGURE 1. The flow of the proposed methodology

authors integrated the use of both k-means clustering and fuzzy c-means clustering to fully make use of each of the techniques' best qualities. The authors made use of k-means clustering for its minimum computation time and, on the other hand, made use of fuzzy c-means' accurate segmentation property. This approach proved to be on par when compared to different state-of-the-art algorithms in segmentation and was also able to output a very good accuracy of the segmented images.

Nisha Pal, in one other paper [12], has made use of thresholding segmentation and watershed segmentation to detect tumor regions in MRI scans of the brain. Even though watershed segmentation is an exemplary technique for analyzing and combining pixels based on their intensity, it has the disadvantage of being very sensitive when dealing with local minima as it creates a watershed for each minimum that it encounters.

III. METHODOLOGY

The approach used in this research work aims to identify and detect tumors from an MRI of a brain. Fig.1. depicts the flow of the proposed methodology. Algorithm 1 describes the methodology.

Algorithm 1:
Step 1: Read the inputted MRI image of the human brain.
Step 2: Convert the MRI image into grayscale.
Step 3: Resize the grayscaled image into a 256*256 image matrix.
Step 4: Remove noise in the image with a median filter.
Step 5: Pass the resultant image obtained in Step 3 to the model.
Step 6: Compute a global threshold for converting an intensity image.
Step 7: Evaluate the morphological operation (Fuzzy C means clustering).
Step 8: Erode the image to get the tumor portion image.

Fig. 2 depicts the proposed system model. In this method, an MRI scan of the brain is chosen by browsing. This selected image is resized to 256*256 and then grayscaled. A gray-scale converted image is generated from a color image using a large matrix with integer values extending from 0 to 255, where 0 and 255 symbolize black and white, respectively. The adaptive median filter technique is then applied to the grayscale image to remove noise. This technique categorizes a pixels as noise by contrasting them with its neighboring image pixels. If a pixel is different from the bulk of its neighbors, it is then replaced by a median pixel. Using the adaptive median filter technique, a smoothed image is obtained[13].

Following noise removal from the input image, the affected area is highlighted with the help of the Balance contrast enhancement technique. This technique compresses the image without altering the grayscale or input image's histogram pattern. It highlights the white pixels, which are the area of interest when compared to the nearest pixels. During this process, contrast images are converted to binary images in order to detect tumorous regions [9].

.

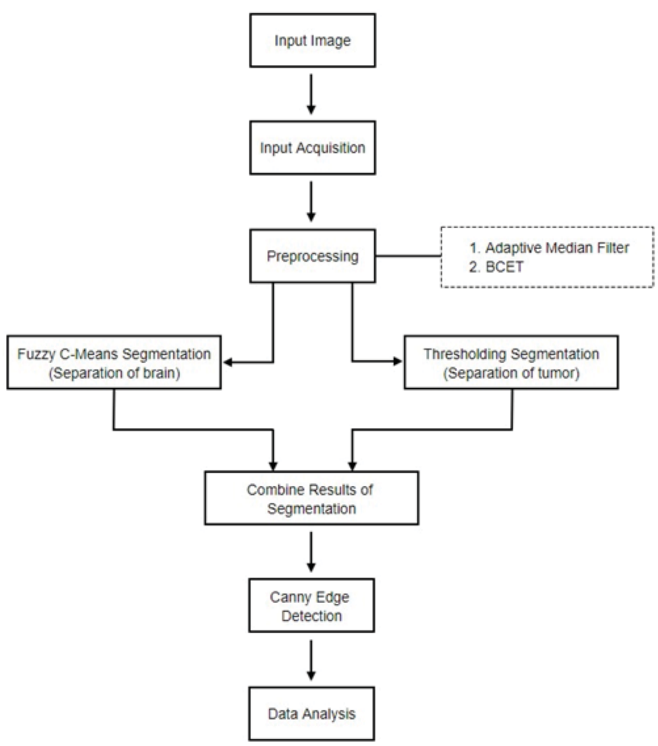

Figure 2. flow chart

Fuzzy C Means Clustering is then performed to identify the segment of the brain region. Based on the distance between the cluster center and the data point, this algorithm assigns membership to each data point that corresponds to each cluster center. The closer the data is to the cluster center, the more likely it is to belong there. Two clusters are assigned in this process, 0 for the brain and 1 for the non-brain (skull) part of the image[14][15]. Then a thresholding technique is used for identifying a segment of the tumor region. In this process, it assigns the threshold value of 256. If the pixel values are more significant than the threshold value, that pixel is converted into white pixels, and the remaining pixels become black. Thresholding is a technique in image segmentation where the pixels of an image are modified to make the analysis of the image simpler. In thresholding, we convert a color or grayscale image into a binary image, i.e., a black and white image. Thresholding is most commonly used in image processing to select areas of interest while ignoring irrelevant parts[16].

The Canny edge detection algorithm is then applied for the edge detection for the combined segmented images [17]. Finally, the presence of white pixels in the tumor-affected region can be identified. Concatenation is then performed between the detected segment region of the tumor image and the standard brain image. Lastly, RGB color was added to the segmentation image, in that red color denotes the tumor region from the initial MRI image.

IV. ANALYSIS

The dataset utilized in the present research is a consolidation of several datasets of MRI images acquired from Kaggle. Experimental analysis was performed to assess the performance of the model. Fig.3. and Fig.4. display the generated edge maps for input MRI scans, working of the methodology mentioned in the previous section. Representing a tumor case and a non-tumor case respectively.

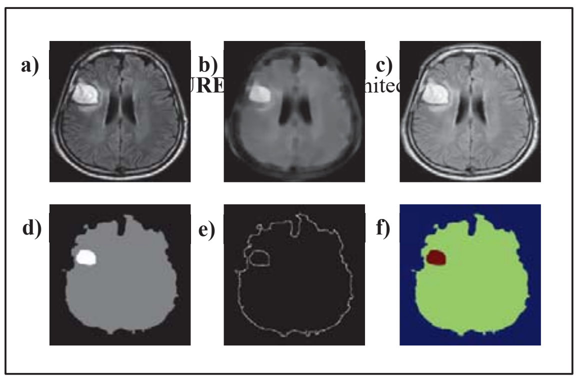

FIGURE 3. a) Resized Image, b) AMF, c) BCET, d) Fuzzy C means, e) Canny Edge Map, f) Color Map

One of the major obstacles with image processing is the testing phase. The testing of image processing tasks is highly resource-consuming [18]. Many complex images are needed as test inputs, and the expected results have to be determined to define the test cases. In the current paper, we have sourced together labelled MRI scans of the brain from multiple sources and used them for the testing phase[19][20][21]. We have gathered random test inputs and tested against the developed project, though the testing is manual. The project has seen up to 300 test inputs, a few of which have been shown in fig. 5. And Table 1 displays the results of the images displayed in fig. 5. Out of the 300 test inputs, about 98% were correctly identified and the results matched with the respective expected outputs.

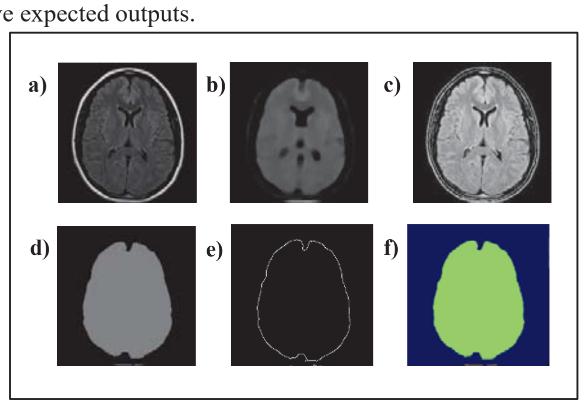

FIGURE 4. a) Resized Image, b) AMF, c) BCET, d) Fuzzy C means, e) Canny Edge Map, f) Color Map

At first, the input image is processed through the filters described in Fig.3 and Fig.4. Following that, various features are extracted implicitly. Every part of the brain tumor must be chosen. Even a tiny portion of the brain tumor is not spared.

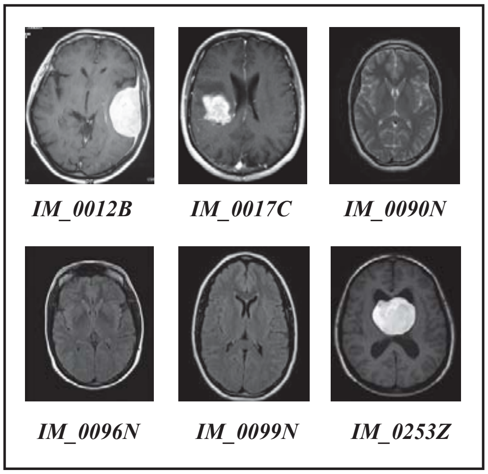

FIGURE 5. Database of test images

TABLE 1. EXPERIMENTAL RESULTS

IMAGE NAME	IMAGE SIZE	RESULT
IM_0099N	409*442	Tumor Absent
IM_0096N	550*664	Tumor Absent
IM_0012B	825*993	Tumor Present
IM_0090N	630*630	Tumor Absent
IM_0017C	400*369	Tumor Present
IM_0253Z	283*338	Tumor Present

the image's contrast. The importance of MRI image pre-processing in enhancing the image's visual effect for subsequent processing cannot be overstated [22]. The images in the dataset are usually of substandard quality, necessitating noise filtering and image sharpening. The acquired image in the dataset is converted into a two-dimensional matrix in a pre-processing step, and the image is transformed into a grayscale image [23]. For the purpose of noise removal, a median filter is used. Following that, the image is enhanced using BCET. The area of affectation (contour) regions in tumors and normal brains can be precisely defined using segmentation. There are various traditional methods for edge detection, but in this research, the Canny edge detection technique was utilized as its edge detection is dependent on neighbor pixels. For testing, a diverse set of MRI images were used.

TABLE 2. PERFORMANCE COMPUTATION

Confusion Matrix	64 (TP)	3 (FN)
	2 (FP)	46 (TN)
Accuracy	0.9565	
Error rate	0.0435	
Sensitivity	0.9697	
Specificity	0.9388	
Precision	0.9552	
F1 score	0.9624	

It can be ascertained that the model built around the proposed methodology produces acceptable results after being evaluated with classification metrics. Table 2 illustrates the model performance results when tested with a dataset sourced from Kaggle [24]. The model makes a successful identification rate of 95.65%, a misidentification rate of 4.35%, sensitivity of 96.97%, specificity of 93.88%, precision of 95.52%, and an F1 score of 96.24%. The Intersection-Over-Union (IoU) is calculated to be 0.927, which is a good score for accurate object segmentation [25]. IoU is an evaluation metric that is generally used to measure the accuracy of an object detector, in this case, the object being the tumorous parts in the input brain scans. The Dice Coefficient score (F1 score) measures to be 0.96, which is indicative of the models effectiveness in segmentation tasks. The Sørensen–Dice coefficient gauges the similarity in data elements; it is used to evaluate segmentation tasks [26].

The numbers from the above section assure the models stable and reliant performance, this proposed model could be scaled up using enterprise hardware to facilitate larger data sets to be analyzed simultaneously.

V. CONCLUSION

In this project, image segmentation technology is used to identify and map the contour of MRI scans of a tumorous brain and a normal brain. The area of (contour) affected regions of tumors and general brains can be defined precisely using segmentation. The results are much more accurate and precise, as shown in the previous section. The accuracy of the final result is determined by how each step is completed. There are several methods to choose from for each step, and the processes that produced the best results were selected. Furthermore, this methodology could be expanded to identify the presence of tumors in other parts of the human body (such as the liver, colon, prostate, breast etc.).

REFERENCES

1. Brain Tumor: Statistics | Cancer.Net
2. Brain Tumors and Brain Cancer | Johns Hopkins Medicine
3. David N Louis, Arie Perry, Pieter Wesseling, Daniel J Brat, Ian A Cree, Dominique Figarella-Branger, Cynthia Hawkins, H K Ng, Stefan M Pfister, Guido Reifenberger, Riccardo Soffietti, Andreas von Deimling, David W Ellison, The 2021 WHO Classification of Tumors of the Central Nervous System: a summary, Neuro-Oncology, Volume 23, Issue 8, August 2021, Pages 1231–1251
4. L. S. G. Kovasznay and H. M. Joseph, "Image Processing," in Proceedings of the IRE, vol. 43, no. 5, pp. 560-570, May 1955, doi: 10.1109/JRPROC.1955.278100.
5. Abubakar, Fari Muhammad. "A study of region-based and contour-based image segmentation." Signal & Image Processing 3.6 (2012): 15.
6. Kumar, Saurabh, et al. "BRAIN TUMOR DETECTION USING IMAGE PROCESSING."
7. Self-organizing Maps (hmc.edu)
8. Jinal A. Shah, et al. “Brain Tumor Detection from MRI Images using Fuzzy C-Means Segmentation” International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 6, June 2016
9. Balance Contrast Enhancement Technique (BCET) | IMAGE PROCESSING (imageeprocessing.com)
10. Sawant, Aaswad, et al. "Brain cancer detection from mri: A machine learning approach (tensorflow)." Brain 5.04 (2018).
11. Eman Abdel-Maksoud, Mohammed Elmogy, Rashid Al-Awadi, “Brain tumor segmentation based on a hybrid clustering technique”, Egyptian Informatics Journal, Volume 16, Issue 1, 2015, Pages 71-81
12. Nisha Pal, “Brain tumor detection using thresholding and watershed algorithm” Journal Of Engineering & Technology, Volume 5, Issue 2, 2019
13. Kesari Verma, Bikesh Kumar Singh, A.S. Thoke, "An Enhancement in Adaptive Median Filter for Edge Preservation", Procedia Computer Science, Volume 48, 2015, Pages 29-36, ISSN 1877-0509
14. Cannon, Robert L., Jitendra V. Dave, and James C. Bezdek. "Efficient implementation of the fuzzy c-means clustering algorithms." IEEE transactions on pattern analysis and machine intelligence 2 (1986): 248-255.
15. Ghosh, Soumi, and Sanjay Kumar Dubey. "Comparative analysis of k-means and fuzzy c-means algorithms." International Journal of Advanced Computer Science and Applications 4.4 (2013).
16. Oliva, Diego, et al. "Multilevel thresholding segmentation based on harmony search optimization." Journal of Applied Mathematics 2013 (2013).
17. Xuan, Li, and Zhang Hong. "An improved canny edge detection algorithm." 2017 8th IEEE international conference on software engineering and service science (ICSESS). IEEE, 2017.
18. Mayer, Johannes and Ralph Guderlei. “On Random Testing of Image Processing Applications.” 2006 Sixth International Conference on Quality Software (QSIC'06) (2006): 85-92.
19. Brain MRI Images for Brain Tumor Detection | Kaggle
20. Brain Tumor MRI Dataset | Kaggle
21. GitHub - guillaumefrd/brain-tumor-mri-dataset: Utilities to download and load an MRI brain tumor dataset with Python, providing 2D slices, tumor masks and tumor classes.
22. Akar, Engin, et al. "Fractal analysis of MR images in patients with Chiari malformation: The importance of preprocessing." Biomedical Signal Processing and Control 31 (2017): 63-70.
23. Image Processing 101 Chapter 1.3: Color Space Conversion | Dynamsoft Document Image Blog
24. Brain Tumor Classification (MRI) | Kaggle
25. Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression (arxiv.org)
26. Zou KH, Warfield SK, Bharatha A, et al. “Statistical validation of image segmentation quality based on a spatial overlap index.” Acad Radiol. 2004; 11(2):178-189.