JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

BRAIN TUMOR SEGMENTATION USING K-MEANS CLUSTERING ALGORITHM

Mrs. Nandhini A[1], Sathiyaseelan M[2]
[1]Assistant Professor (SG), Department of Computer Applications, Nehru College of Management, Coimbatore, TamilNadu, India.
[2] II MCA, Department of Computer Applications, Nehru College of Management, Coimbatore, TamilNadu, India.

Abstract: *An abnormal and uncontrollable development of brain cells is known as a brain tumor. Because of the problems with the brain's anatomy, diagnosing a brain tumor can be difficult. These magnetic resonance images that were gathered as big data can be used to identify brain cancers. We can detect a variety of disorders and investigate the development of the human brain using the rich anatomical information obtained from these MR images. Because of the large image library, brain tumor identification becomes increasingly difficult. Therefore, an algorithmic method is needed to provide a more rapid and accurate clinical diagnosis. In order to precisely diagnose the area of brain tumor, the primary focus of this work is the brain segmentation of MR images using the k-means clustering algorithm. The brain tumor is found following segmentation, which is done using the k-means clustering algorithm.*

Keywords: Brain tumor, detection and segmentation, magnetic resonance imaging (MRI), k-means clustering algorithm.

I. Introduction:

A brain tumor is a mass, or lump, that develops in the brain as a result of unchecked brain cell division and growth. Our bodies use growth, division, and multiplication of cells to support normal functions like cell replacement and damage repair. However, errors can occasionally occur during cell division and growth. And aberrant cells start to form.

These aberrant cells are usually eliminated by the body's defensive mechanisms, but on rare occasions, they proliferated and aggregate into a mass of cells. We refer to this as a tumor. A primary brain tumor forms in the brain when this occurs. The brain contains approximately 100 billion nerves that are arranged in overlapping pattern. Therefore, identifying the locations of the brain tumor is a difficult task. Individuals of various ages, including teenagers, may experience the symptoms of a brain tumor. Individuals under the age of twenty-one accounted for roughly one-third of all new brain cancer diagnoses in recent years, while individuals between the ages of twenty and thirty-four accounted for another 9%. The brain's unchecked cell proliferation is the cause of the tumor. Primary brain tumors can be classified as either benign or malignant. A brain tumor has clearly defined boundaries, grows slowly, and is tiny. It can be totally eliminated through surgery, and it does not spread to other parts of the body or the spinal cord. The malignant kind of tumor grows quickly, damages normal brain cells, and has the potential to spread to others brain or spinal cord regions. It is dangerous and might not better.

II. Related works:

The k-mean clustering technique is a popular algorithm in picture segmentation system. Researchers have made multiple attempts to increase the k-means algorithm's efficacy and efficiency.

Stephen suggested a technique for selecting K randomly selected examples as seeds from a database. The computational difficulty of this method is its drawback; the k-means algorithm must be run multiple times before each instance can be assigned, which can be very taxing on a big database.

Douglas and Michale proposed splitting the dataset into chunks and using k-means on each block to choose an acceptable starting solution. Even though the aforementioned algorithms can be some what helpful in locating suitable initial centers, they are rather sophisticated, and some of them incorporate the k-means algorithm. As a result, cluster initialization must be still be done using the random technique.

Shapiro and haralick there aren't complete theory of picture segmentation since there isn't complete theory of clustering. The following qualitative guidelines for effective picture segmentation have been created by them.

This indicates that the methods used for segmenting images are typically and hoc and vary in how they highlight one or more of the desired attributes. Consequently, each images segmentation algorithms ultimate implement is heavily dependent on the application's outcomes.

Nonetheless, these variations typically revolve on the selection of parameters or approaches for customizing particular aspects of the image.

III. Methodology:

Image acquisition is the first step in the process of using MR scans to identify brain tumors. Since we are utilizing an already-existing dataset, the suggested methodology does not include an image capture step. Data collecting, picture preprocessing, data augmentation, feature extraction, feature selection, and class identification make up our methodology. The diagrammatic representation of figure 1 is displayed below.

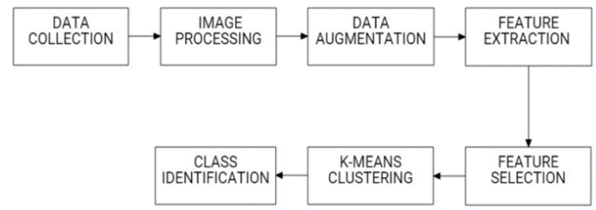

Figure1. Proposed methodology

Data collection:

For each T1, T1c, T2 and flair cases, the BRATS dataset contains a three-dimensional magnetic resonance picture of the brain with four distinct

Modalities. Doctors assembled this dataset, which contained three tumor sub-regions, after closely observing the cases, These observations

were integrated into three-sub regions: enhancing tumor, complete tumor, and tumor care. The enhancing tumor, peritumoral edema, and non-enhancing tumor were identified.

Image processing:

To ensure accurate result creation, excess noise is eliminated from all images before they are utilized for categorization. Additionally, pre-processing techniques such as mean normalization, standardization, noise reduction, picture filtering, scaling, and geometric modifications are used to achieve image pre-processing. The features found in specific photos are advanced through the use of image processing. Since every outcome in computer vision depends on its features, image processing is mostly utilized in this field. There are theories circulating that pre-processing an image may alter its genuine characteristics.

Data augmentation:

When analyzing data, the process known as "data augmentation" is used to add new aggregated data that is derived from existing data or slightly modified copies of the original data. In the process of training machine learning models, this serves as a regularization and helps to minimize overfitting. This and oversampling in data analysis are closely related. Similar to implicit regularization, data augmentation is a widely used method for enhancing the generality of deep neural networks. When there is a lack of trustworthy, high-quality data and creating new examples takes money and effort, it becomes crucial. When analyzing medical photos, this is a fairly prevalent issue, particularly when finding representative tumors. Magnetic resonance imaging approaches for brain tumors with enhanced data. In data analysis, strategies such as down sampling and oversampling are employed to modify the distribution of classes within a data set. The phrase is used in machine learning as well as statistical sampling and survey design techniques. The opposite approaches of approximation are down sampling and oversampling. There are even more straightforward oversampling methods that use algorithms like composite prime sampling to create false data points.

Feature extraction:

Feature extraction helps in pattern classification by extracting important shape information from the pattern. Pattern recognition and image processing involve feature extraction, a kind of dimensionality reduction. Taking the most significant information out of the original data and expressing it in a lower-dimensional space is the main goal of feature extraction. A reduced representation set of features, often referred to as a feature vector, is created from input data that is too large for analysis and is thought to be redundant. The process of turning unprocessed data into a collection of features is called feature extraction.

Future selection:

We are unable to categorize using all the features that are extracted, regardless of how many or how few there are. As a result, a few crucial features are chosen, and the outcome is assessed based on those aspects. In addition, after fewer features are included, the model performance is immediately impacted and the processing cost of implementing the model is also decreased. Certain features are especially useful because our model relies on k-means clustering, which is an unsupervised learning technique. There are several features selection techniques accessible for supervised learning, including filter, wrapper, and embedding techniques.

K-means clustering:

The most popular algorithm for the unlabeled dataset is K-which clustering. It is one of the algorithms for unsupervised learning that addresses clustering issues. K-means operates by creating multiple clusters based on specific criteria, with a predetermined number of clusters (k). Since it uses an iterative process to create k distinct clusters based on comparable qualities, it is also referred to as an iterative algorithm. Additionally, as that method links each cluster to a specific centroid, it is sometimes referred to as a centroid-based algorithm.

A straightforward illustration that is clarified by the graphic below might be taken into consideration. In figure 2, distinct data points are displayed. However, if a specific feature is unique to that type of data, it can be designated as the feature for gathering all other similar types. Similarly, in figure 3, data points are distinguished based on color.

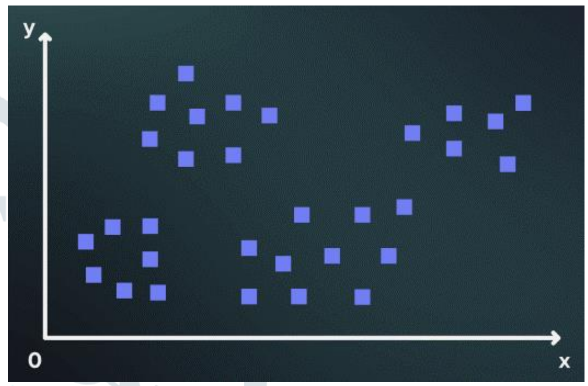

Figure 2. before k-means clustering

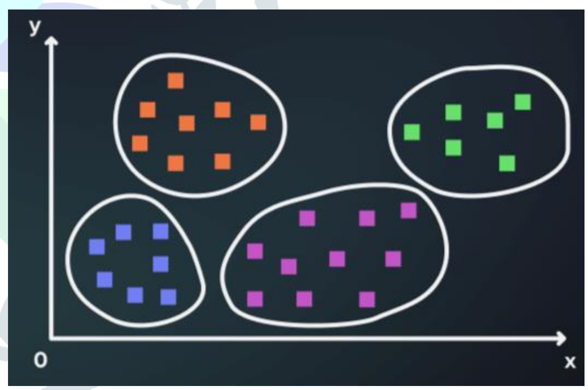

Figure 3. after k-means clustering

Class identification:

Brain tumor can be differentiated into primary and secondary types. Primary brain tumor begins and spread within the brain. Many of the first brain tumors are not harmful. Metastatic brain tumors, commonly referred to as a secondary brain tumor, develop when cancer cells spread to the brain from the other organs like the chest or lungs.

Primary brain tumor: Brain tumors that originate directly inside the brain or CNS are classified as either malignant or benign.

Metastatic brain tumor: Cancer cells from other parts of body that have metastasized to the brain are the source of metastatic brain tumor.

IV. Results:

Using our suggested algorithm, the location of brain tumors in brain MR images is determined. A portion of the brain MR images with tumors that were acquired to test by algorithm.

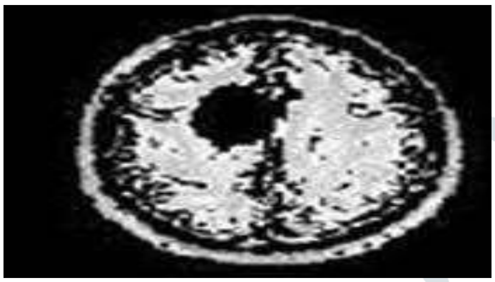

Figure 4. Applying proposed algorithm

Figure 5. clustering image

Figure 6. tumor detection

V. Conclusion:

Medical professionals must segment brain images in order to plan and perform surgery. With the use of the k-means clustering technique, we have developed a computer-aided system for segmenting brain MR images and locating tumors. Image capture, pre-processing, and k-means clustering are the three stages of the suggested brain tumor detection method using our database's various brain MR pictures, we were able to segment the tumor.

References

[1] Dr. A. Geetha, A. Mary Ezhilvani "A CLASSIFICATION TECHNIQUES SURVEY USED FOR BRAIN TUMOR DETECTION" © 2020 IJCRT | February 2020 | Volume 8, Issue 2 | ISSN: 2320-2882.

[2] Instructor A.S. Bhide1 "BRAIN SEGMENTATION USING FUZZY C MEANS CLUSTERING TO DETECT TUMOR REGION" Priyanka Patil2, Shraddha Dhande3, ISSN: 2277-9043 Volume 1, Issue 2, April 2012, International Journal of Advanced Research in Computer Science and Electronics Engineering.

[3] Vidya Zope2, Bhavika Mahadik1, Tushar Jumani1, Rahul Godhani1, Gunjan Gurbani1, and Tushar Jumani "A High Impact Factor, Monthly, Peer-Reviewed Journal on the Comparison of Supervised and Unsupervised Learning Algorithms to Brain Tumor Detection" published in the International Journal of Innovative Research in Science, Engineering, and Technology Visit: January 2019, Vol. 8, Issue 1.

[4] Dr. Nadira Banu, Malathi R The article "Brain Tumor Identification and Detection Using K-Means Clustering Technique" was published in a special issue of the International Journal of Advanced Networking and Applications (IJANA) by Kamal A. R..

[5] Systematic image processing for brain tumor diagnosis: A Type-II fuzzy expert system approach, Applied soft computing, M.H. Fazel Zarandia, M. Zarinbala, and M. Izadi b. (2011) 11,285-294

[6] A. Suman Tati Raju, Symposium of Discrete Algorithms, Image Segmentation using k-means clustering, EM and Nominalized Cuts, July 2008. (2010) Fast Image Segmentation Based on K-Means Clustering with Histograms in HSV Color Space by Tse-Wei Chen, Yi-Ling Chen, and Shao-Yi Chien, Journal of Scientific Research, ISSN I452-2I6X, Vol. 44, No. 2, pp. 337–35

[7] In the First Workshop on High Performance Data Mining, K. Alsabti, S. Ranka, and V. Singh (Mar. 1998) presented An

Efficient k-means Clustering Algorithm.
[8] Mohamed Sathik, S. Zulaikha Beevi M (20I0). European Journal of Scientific Research, Vol. 41, No. 3, pp. 437–45, "An Effective Approach for Segmentation of MRI Images: Combining Spatial Information with Fuzzy C-Means Clustering."

[9] Color Skin Segmentation Using K-Means Cluster, K.S. Ravichandran and B. Ananthi (2009), International Journal of Computational and Applied Mathematics, Volume 4 Number 2, pp.153-157

[10] Douglas Steinley and Michael J. Brusco (2007) Journal of Classification DOI: 10.1007/s00357-007-0003-0,24:99-121 Initializing K-means Batch Clustering: A Critical Evaluation of Several Techniques

[11] In IEEE Transactions on Fuzzy Systems, D. Van De Ville, M. Nachtegael, D. Van Der Weken, E.E. Kerre, W. Philips, and I. LeMahieu (2003) discussed noise reduction by fuzzy picture filtering. Pfefferkorn Conference on Signal and Image Processing in Microscopy and Microanalysis, Proceedings, Segmentation (September 1991), pp. 299-314.

[12] A unified patch-based technique for brain tumor diagnosis using fusion features; M. Sharif, J. Amin, M. W. Nisar, M. A. Anjum, N. Muhammad, and S. Ali Shad, Cogn. Syst. Res., vol. 59, pp. 273–286, 2020, Doi:10.1016/j.cogsys.2019.10.001.

[13]I.Processing(pp.3984)"a71E05Cfc49Ab0777B82Ca94D181 F779149F."

[14] In J. Vis. Exp., no. 64, June 2012, N. Jeremy Hill et al. recorded human electrocorticographic (ECoG) signals for neuroscientific study and real-time functional cortex mapping; doi: 10.3791/3993.

[15] "The brain tumor microenvironment," N. A. Charles, E. C. Holland, R. Gilbertson, R. Glass, and H. Kettenmann, Glia, vol. 59, no. 8, pp. 1169–1180, 2011, doi: 10.1002/glia.21136

[16] The article "The global k-means clustering algorithm" was published in 2003 by A. Likas, N. Vlassis, and J. J. Verbeek in Pattern Recognition. It may be accessed at https://doi.org/10.1016/S0031-3203(02)00060-2.

[17] The article "Cluster center initialization algorithm for K-means clustering" was published in 2004 by S. S. Khan and A. Ahmad in Pattern Recognition Letters, vol. 25, no. 11, pp. 1293–1302, doi: 10.1016/j.patrec.2004.04.007.