Asian Journal of Computer Science and Technology
ISSN: 2249-0701 (P) Vol.8 No.3, 2019, pp.23-26

DOI: https://doi.org/10.51983/ajcst-2019.8.3.2736

Big Data Analytics for Brain Tumour Detection Using Subspace Clustering

Padmavathi Vanka[1] and T. Sudha[2]
[1]Research Scholar and [2]Professor, [1&2]Department of Computer Science
Sri Padmavathi Mahila Viswa Vidyalayam, Tirupati, Andhra Pradesh, India
E-Mail: vanka.padmavathi@gmail.com, thatimakula_sudha@yahoo.com
(Received 23 July 2019; Revised 7 August 2019; Accepted 4 September 2019; Available online 15 September 2019)

***Abstract* - In healthcare industry data mining holds a great prospective to enable health systems steadily use big data and analytics to identify inefficiencies which improves care and reduces cost. Due to the slower rate of technological adoption in health care, software industry lags in implementing effective data mining and analytic strategies. The segmentation, detection, and extraction of infected area of tumours from magnetic resonance images (MRI) are a primary distress but a tiresome task performed by radiologists or clinical experts. To improve the performance and reduce the complexities in segmenting the MRI data new methods need to be introduced. A 3D magnetic resonance images can give better accuracy in detecting the tumour area. In this paper we proposed subspace clustering techniques to detect brain tumours using 3D images.**
***Keywords*: Subspace Clustering, Slicer Tool, CLIQUE**

I. INTRODUCTION

Image segmentation refers to the process of partitioning a digital image into multiple regions. The goal of segmentation is to change the representation of an image to be more meaningful and easier to analyze. It is used in order to locate objects and boundaries in images. The result of image segmentation occurs as a set of regions that collectively covers the entire image [1]. Therefore, medical image segmentation plays a significant role in clinical diagnosis. It can be considered as a difficult problem because medical images commonly have poor contrasts, different types of noise, and missing or diffusive boundaries. The anatomy of the brain can be scanned by Magnetic Resonance Imaging (MRI) scan or computed tomography (CT) scan. The MRI scan is more comfortable than CT scan for diagnosis. It is not affect the human body because it does not use any radiation. It is based on the magnetic field and radio waves [2]. On the other hand, brain tumour is one of the leading causes of death among people. It is evidence that the chance of survival can be increased if the tumour is detected correctly at its early stage. In most cases, the physician gives the treatment for the strokes rather than the treatment for the tumour. Therefore, detection of the tumour is essential for the treatment. The lifetime of the person who affected by the brain tumour will increase if it is detected early [3]. Thus, there is a need for an efficient medical image segmentation method with some preferred properties such as minimum user interaction, fast computation, accurate, and robust segmentation results [4].On the other hand, image segmentation algorithms are based on one of the two fundamental properties of image intensity values: discontinuity and similarity In the formal category, the segmentation approach is based on partitioning the processed image based on changes in intensity, such as edges and corners. The second one is based on partitioning an image into regions that are similar due to a set of pre-defined criteria. Therefore, there are many segmentation techniques which can be broadly used, such as histogram based methods, edge-based methods, artificial neural network based segmentation methods, physical model based approaches, region-based methods (region splitting, growing, and merging), and clustering methods (Fuzzy C-means clustering, K-means clustering, Mean Shift, and Expectation Maximization).

There are many challenging issues to image segmentation like development of a unified approach that can be applied to all types of images and applications [10]. Even, the selection of an appropriate technique for a particular kind of image is a difficult problem. Thus, there is no universal accepted method for image segmentation. So, it remains a challenging problem in image processing and computer vision fields. One view of image segmentation is a clustering problem that concerns how to determine which pixels in an image belong together most appropriately. Cluster analysis serves as a pre-processing step for other algorithms, such as classification that would then operate on detected clusters. We used image segmentation techniques based on clustering to detect the brain tumour and calculating the tumour area. Subspace clustering is an extension of traditional clustering that seeks to find clusters in different subspaces within a dataset [11]. Often in high dimensional data, many dimensions are irrelevant and can mask existing clusters in noisy data. Feature selection removes irrelevant and redundant dimensions by analyzing the entire dataset. Subspace clustering algorithms localize the search for relevant dimensions allowing them to find clusters that exist in multiple, possibly overlapping subspaces. There are two major branches of subspace clustering based on their search strategy. Top- down algorithms find an initial clustering in the full set of dimensions and evaluate the subspaces of each cluster, iteratively improving the results [12]. Bottom-up approaches find dense regions in low dimensional spaces and combine them to form clusters. Often, bottom-up algorithms are able to find clusters of various shapes and sizes since the clusters are formed from various cells in a grid. This means that the clusters can overlap each other with one instance having the potential to be in more than one cluster. It is also possible for an instance to be considered an outlier and not belong any cluster. Subspace K means clustering algorithm can detect a brain tumour faster than Fuzzy C-means and K means algorithms in 3D images.

II. RELATED WORK

Medial image segmentation is a burning research area where several researchers have suggested various methodologies and algorithms for image segmentation. Priyanka Shah, Manila Jeshnani, Sagar Kukreja and Priyanka Ailani [5] proposed various segmentation methods and the algorithms that can be used for segmenting the MRI images. The K means algorithm is deterministic nature, which makes it independent of cluster initialization, and the ability to identify nonlinearly separable clusters in input space. Fuzzy C Means algorithm is powerful unsupervised method for the analysis of data and construction of models but not efficient time consuming [6]. Watershed Segmentation is efficient and accurate, in this the process is divided in 2 parts, hence making it easier to use [7]. But it gives over segmentation, which can be solved using topological gradient. The selection of the algorithm will depend on the need of the user. K-means clustering algorithm can be used to classify the objects in an image. Watershed segmentation can be used for the tumour regions that have high intensity values. E. Abdel-Maksoud, M. Elmogy, and R. Al-AwadiImage segmentation plays a significant role in medical image [8]. In the field of medical diagnosis, an extensive diversity of imaging techniques is available presently, such as CT and MRI. MRI is the most effectively image model used for diagnostic image examination for brain tumour. Original Fuzzy C-means algorithm fails to segment image corrupted by noise, outliers, and other imaging artifacts. Therefore, we developed a new approach that integrates the K-means clustering algorithm with the Fuzzy C-means algorithm to detect brain tumour accurately and in minimal execution time. Ramish B. Kawadiwale and Milind E. Rane proposed an accurate detection of brain tumour using clustering techniques [9]. MR brain images were taken and processed through clustering techniques thus giving efficient end results for detection of tumours. These techniques give efficient and effective results as compared to previous researches. These techniques were applied on various images and results were extraordinary. Proposed research is easy to execute, scalable and can be modified easily. ShravanRao, Meet Parikh, Mohit Parikh, Chinmay Nemade proposed the system aims to challenge a very alarming issue by providing different results and comparing them with modified techniques [10]. Fuzzy-C means achieves better clustering due to membership concept, however time complexity of the system is high. When considering system accuracy with respect to ground truth, Adaptive-K clustering renders better segmentation accuracy and tumour extraction.

R. Agrawal, J. Gehrke, D. Gunopulos, and P. Raghava proposed high dimensional data is increasingly common in many fields [12]. Subspace clustering attempts to integrate feature evaluation and clustering in order to find clusters in different subspaces. Bottom-up algorithms integrate the clustering and subspace selection by first selecting a subspace, then evaluating the instances in the context.

III. PROPOSED METHOD

The proposed method consists of image acquisition, pre-processing, segmentation, feature extraction and result validation. Pre-processing is done by median or Gaussian filtering methods where the noise in the image is removed with default structuring. After eliminating the noise the image is segmented using subspace clustering and Fuzzy C-Means clustering algorithms. Feature extraction is done by thresholding along with region growing and level set contouring. CLIQUE [13] which is a bottom up clustering technique was one of the best algorithms proposed that attempted to find clusters within subspaces of the dataset. As described above, the algorithm combines density and grid based clustering and uses an APRIORI style technique to find cluster subspaces. Once the dense subspaces are found they are sorted by coverage, where coverage is defined as the fraction of the dataset covered by the dense units in the subspace. Finally area of tumour is auto-calculated along with validation with respect to ground truth images and accuracy of the system is calculated.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as colour, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristics. When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create 3D reconstructions with the help of interpolation algorithms like marching cubes.

The K-means algorithm is an iterative technique that is used to partition an image into *K* clusters. The 3D image is sliced into 2D images and performs K-means clustering and finally all the subspaces. The basic algorithm is

1. Pick *K* cluster centres, either randomly or based on some heuristic method, for example K-means++.
2. Assign each pixel in the image to the cluster that minimizes the distance between the pixel and the cluster center.
3. Re-compute the cluster centres by averaging all of the pixels in the cluster.
4. Repeat steps 2 and 3 until convergence is attained (i.e. no pixels change clusters).

In this case, distance is the squared or absolute difference between a pixel and a cluster center. The difference is typically based on pixel color, intensity, texture, and location, or a weighted combination of these factors. *K* number of clusters can be selected manually, randomly, or by a heuristic. This algorithm is guaranteed to converge, but it may not return the optimal solution. The quality of the solution depends on the initial set of clusters and the value of *K*.

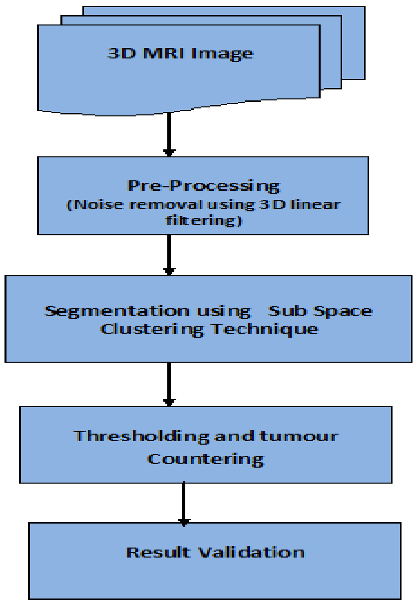

Fig. 1 Frame work for brain tumour detection

IV. IMPLEMENTATION

3D Slicer is free open source software (BSD-style license) that is a flexible, modular platform for image analysis and visualization. 3D Slicer provides image registration, processing of DTI (diffusion tractography), an interface to external devices for image guidance support, and GPU-enabled volume rendering, among other capabilities. 3D Slicer has a modular organization that allows the addition of new functionality and provides a number of generic features not available in competing tools.

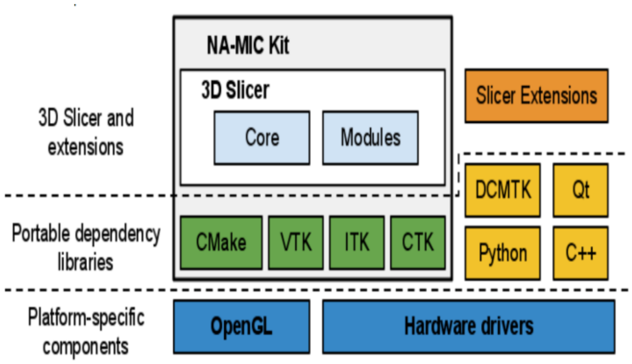

Fig. 2 3D slicer echo system

The license has no restrictions on use of the software in academic or commercial projects. Slicer's platform provides functionalities for segmentation, registration and three-dimensional visualization of multimodal image data, as well as advanced image analysis algorithms for diffusion tensor imaging, functional magnetic resonance imaging and image-guided radiation therapy. Standard image file formats are supported, and the application integrates interface capabilities to biomedical research software.

Fig. 3 Images of a tumour patient in 3D slicer

The radiological images produced in dicom format can be converted into nrrd file format to upload in 3D slicer. Click on load data in the slicer and then choose the dataset to add. Move to visualization tool to view image.

V. CONCLUSION

Image segmentation plays a significant role in medical image. In the field of medical diagnosis, an extensive diversity of imaging techniques is available presently, such as CT and MRI. MRI is the most effective image model used for diagnostic image examination for brain tumour. The MRI scan is more comfortable than CT scan for diagnosis. The 3D evaluation of the brain tumour detection using 3D slicer can be carried out which accurately finds the brain tumours. As well as to increase the efficiency of the segmentation process, an intensity adjustment process will provide more challenging and may allow us to refine our segmentation techniques to the MRI brain tumour segmentation.

REFERENCES

[1] V. Janani, and P. Meena, “Image segmentation for tumour detection using fuzzy inference system”, *International Journal of Computer Science and Mobile Computing (IJCSMC)*, Vol. 02, No. pp. 244–248, 2013.

[2] J. Patel and K. Doshi “A study of segmentation methods for detection of tumour in brain MRI”, *International Journal of advanced electronics and electrical engineering*, Vol. 04, No. pp. 279–284, 2014.

[3] M. Rohit, S. Kabade and M.S. Gaikwad, “Segmentation of brain tumour and its area calculation in brain MRI images using K-mean clustering and Fuzzy C-mean algorithm”, *International Journal of Computer Science Engineering and Technology (IJCSET),* Vol. 04, No. pp.524–531, 2013.

[4] H.A. Aslam, T. Ramashri, and M.I.A. Ahsan, “A new approach to image segmentation for brain tumour detection using pillar K-means

algorithm", *International Journal of Advanced Research Compution and Communication Engineering*, Vol. 2, pp. 1429–1436, 2013.

[5] Priyanka Shah, Manila Jeshnani, SagarKukreja and PriyankaAilani, "Survey on Algorithms for Brain Tumour Detection", *International Journal of Computer Science and Information Technologies*, Vol. 08, No. pp. 56–58, 2017.

[6] NileshBhaskarraoBahadure, Arun Kumar Ray and Har Pal Thethi, "Image Analysis for MRI Based Brain Tumour Detection and Feature Extraction Using Biologically Inspired BWT and SVM", *International Journal of Biomedical Imaging*, Vol. 2017, Mar. 2017.

[7] S. Damodharan and D. Raghavan, "Combining tissue segmentation and neural network for brain tumour detection", *InternationalArab Journal of Information Technology*, Vol. 12, No. pp. 42–52, 2015.

[8] E. AbdelMaksoud, M. Elmogy, and R. Al-Awadi, "Brain tumour segmentation based on a hybrid clustering technique," *Egyptian Informatics Journal*, Vol. 16, No. pp. 71–81, 2014.

[9] B. Ramish. Kawadiwale and Milind E. Rane, "Clustering Techniques for Brain Tumour Detection", *Association of Computer Electronics and Electrical Engineers*, 2014.

[10] ShravanRao, Meet Parikh, Mohit Parikh and Chinmay Nemade "Implementation of Clustering Techniques For Brain Tumour Detection", *International Research Journal of Engineering and Technology*, Vol. 03. No. pp. 517–521, Apr. 2016.

[11] Lance Parsons, Ehtesham Haque, and Huan Liu "Subspace Clustering for High Dimensional Data: A Review", *Sigkdd Explorations*, Vol. 6, pp.90–105, 2004.

[12] R. Agrawal, J. Gehrke, D. Gunopulos, and P. Raghava, "Automatic subspace clustering of high dimensional data for data mining applications", *In Proceedings of the 1998 ACM SIGMOD International conference on Management of data, ACM Press*, pp. 94–105, 1998.

[13] C.H. Cheng, A.W. Fu, and Y. Zhang, "Entropy-based subspace clustering for mining numerical data", *In Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM Press*, pp. 84–93, 1999.