



# An Efficient Him Technique for Tumour Detection From MRI Images

Deepak Kokate, Jijo Nair

<sup>1</sup>M. Tech. Research Scholar, <sup>2</sup>M. Tech. Research Guide,  
<sup>1,2</sup>Dept of CSE Oriental College, Bhopa, Madhya Pradesh, India

## ABSTRACT

Data mining techniques are widely used for data processing from large data set such as data center and data warehouse. An Image mining technique is a new form of data mining technique in the processing of image data. In the medical field, day by day size of medical images data is increasing. MRI images are one of them. The medical images like as CT scan, MR images are widely used in brain tumor detection, cancer detection from the human body. It is quite challenging and complicated work to detect abnormal cells and tissue such as tumor from MR image data sets. Due to higher importance and demand of medical image data, it is necessary to process it correctly and efficiently. Image Segmentation has an important role in the field medical image processing. In that way, MRI has become a useful medical diagnostic tool for the diagnosis of brain & other medical images.

In this paper, we are presenting a new hybrid image mining technique (HIMT) for MRI Image processing. The proposed HIMT uses combined strategy of clustering method Fuzzy C-Mean with the Genetic algorithm and SVM classifier. The main key feature of proposed method is it can able assigns and processed two or more than two clusters as compared to K-Means method where data point must exclusively belong to one cluster center and genetic algorithm is used as and optimization tool which helps to achieve results in less time. Proposed HIMT and existing method (K- Means clustering method with GA) both are implemented over MATLAB tool and various performance measurement parameters such as detection rate, area or size and time are calculated. Simulation results are clearly influenced that proposed HIMT method performs outstanding over existing method.

**Keywords:** Data Mining, Image mining, MRI Images, K-means Clustering, C-Mean Clustering, Genetic Algorithm and HIMT

## 1. INTRODUCTION

Data mining provides knowledge discovery from large data sets such as data warehousing and data center. Knowledge discovery from the large data set (alphanumeric databases like as relational databases) has been done by data mining. Due to new advances and improvement of secondary and tertiary level storage capacity, which mainly coupled with a low-level storage cost, and also based on a non-standard data sets (such as data in form of pictures and images) is being increased and accumulated. This continues huge collection of data in form of images also needs to be attention to discover knowledge and information. The main issue with an image mining is its slow process.

Various combine strategy such as a content-based image retrieval (or CBIR) with other steps such as image capturing, storing, understanding and analysis make it's more easier. Various works have been done by researchers in image mining field but still, an efficient scheme is required to knowledge from a set of images from a database perspective. This research work mainly focused on efficient image mining scheme for MR images to discovered knowledge. This research work also covers various unexplored area that can covers various benefits in the field of image processing and mining of medical MR images.

### 1.1 NEED OF IMAGE MINING-

An Image is a collection of pixels. Image mining mainly refers to the extract all the implicit knowledge, image data relationship, or other designs not explicitly stored inside images. Image mining is over just an

extendable of existing data mining technique to image sector.

## 1.2 APPLICATIONS OF IMAGE MINING-

Image mining are used in various fields. Different applications of image are-

- In field of medical for diagnosing diseases.(such as Brain tumour)
- In field of Space Research field.
- In field of Satellite cloud Imagery (Such as detecting, copying, an unauthorized image on the internet)
- In field of Remote sensing field.
- In field of wild plant detection (such as egeria detection).
- In field of Natural scene recognition
- In field of Agriculture.
- In field of educational.
- In field of industrial work.

## 1.3 MRI IMAGES-

A MRI or Magnetic resonance image is technique which is used to capture a high-resolution image from inside of the human body areas such as brain. A MR image is basically a type of NMR or nuclear magnetic resonance. This NMR is widely used by researchers, chemists in study of molecules and their properties.

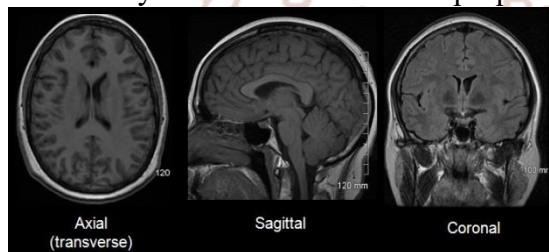


Figure 1.3 MRI Images

The technique is called magnetic resonance imaging or MRI instead of nuclear magnetic resonance image or NMR, due to the reason of (PNA) public's negative associations in late year 1970 [5].

## 2. EXISTING METHODS

Following methods are widely used by various researchers in area of tumor detection and analysis for brain MRI images-

### 2.1 K-MEANS CLUSTERING-

A clustering is a technique which basically divides a collection of data into a small number of areas or groups called as cluster. It's one of the popular methods is k-means clustering. In clustering by K-Means method, a set of data are partitions into various small k number. This method mainly classifies a data

set into various disjoint clusters denotes by K. A K-Means clustering method mainly consists two phases. First phase mainly calculates the various k-centroids. Another phase (Second phase) of K-Means selects each point of the cluster which is mainly near to the centroid from the particular data point.

**Advantages-** Clustering by using K-Means have several advantages such as its easy in implement, simple and best for small set.

**Limitations-** The main limitation of K Mean clustering is the quality and type of a final clusters results are totally depends on the type and quality of initial centroid selection. Means if the selections of initial centroid are based on random selection concept than it will generates different outcome results for an initial center. So in this method the initial selection of center should be very carefully by that we can get our desired segmentation.

### 2.2 IMAGE SEGMENTATION BY WATERSHED ALGORITHM-

The Watershed Transform method uses a unique concept for data and image segmentation. This method uses a new type of area or region growing concept based on an efficient image gradient for processing of various digital images. The main concept behind the Watershed Transform is its visualization process for an image. It uses total three dimensions, in which two are spatial coordinates based and another is based on gray levels.

**Advantages-** Watershed method provides following advantages over various previously developed image segmentation schemes -

- Watershed method provides closed and connected regions boundaries in result over existing methods. Various existing edge based methods generates a disconnected boundaries results which mainly required post-processing steps to generate a closed regions.
- The boundaries generates by Watershed method as resulting regions are correspond to a contours which mainly appear in the image as obvious contours of objects. This is in contrast to divide and merge algorithm. The first method divides or split the contents and second is finally use to merge them.
- In Watershed method the union or sum of all the connected areas or regions are always connected with entire regions.

### 2.3 GENETIC ALGORITHM-

In a Genetic Algorithm, a population of strings called chromosomes which encode candidate solutions to an optimization problem evolves toward better solutions. In GA method the evolution process mainly starts from selection of a population from a randomly generated individual values. In each population generation, only the fitness value is selected. From each of the individual stochastically are selected from the total current population (which is based on their fitness values), and it modified by using GA operators such as crossover and mutation, to or form a new population value.

### 2.4 FUZZY C-MEAN CLUSTERING

Fuzzy C-means is a clustering method which mainly allows various pieces of data or items to relate to two or more than one cluster. This is frequently used in the fields of object recognition, computer vision, and medical image processing. The segmentation results are obtained by using fuzzy classification such as FCM [4].

Unlike the k-means which are a hard classification method mainly allows a data object to strictly belong to the only single cluster. FCM allows a data object or pixel set to relate to various multiple class set without varying their degree of memberships.

A FCM algorithm is quite effective for image segmentation. A fuzzy C-Means clustering method also called FCM [1,3] was firstly introduced by researcher Dunn in year 1974 and later it was extended by researcher Bezdek in year 1981. This algorithm is mainly an iterative data clustering method, which produces an efficient optimal C-partition by minimizing the various weights within the sum of squared error objective function.

### 3. PROBLEM STATEMENT

An image mining is a vast area for research. Image mining methods are widely using by researchers in processing of medical based images such as MRI. Segmentation of MRI images encounters with several issues. A MR image mainly exists in the homogeneous aspects of image pixels in various imaging areas. It makes quite difficult to gather correct and accurate information from MRI data set.

A MR image is basically contents abnormal and normal area captured for brain or any other body internal parts. MR image contents pixel sets for soft

tissue like as brain tissue or a liver tissue. For MR images, any homogeneous requirement or aspects of pixels makes the image segmentation method fail or unsuccessful with the entire MRI image, so the result for one image might be different from another image.

Another difficulties or challenge in existing image processing and mining method is its poor detection quality, poor ratio and detection time for area. Existing image mining researchers were investigates their research by using existing frameworks and method. The following problems are faced by the existing system-

- Poor detection rate
- Poor detection time
- Results are not accurate
- Growth rate and volume are not correct

### 4. PROPOSED METHOD

In this image mining research work, we are presenting a new hybrid image mining (HIM) model for MRI Image processing. The proposed HIMT method uses combined strategy of Fuzzy C-Mean clustering, SVM classifier with Genetic algorithm. Proposed method uses total three phases for tumour detection from various MRI images (figure 4.1) describes the complete phases-

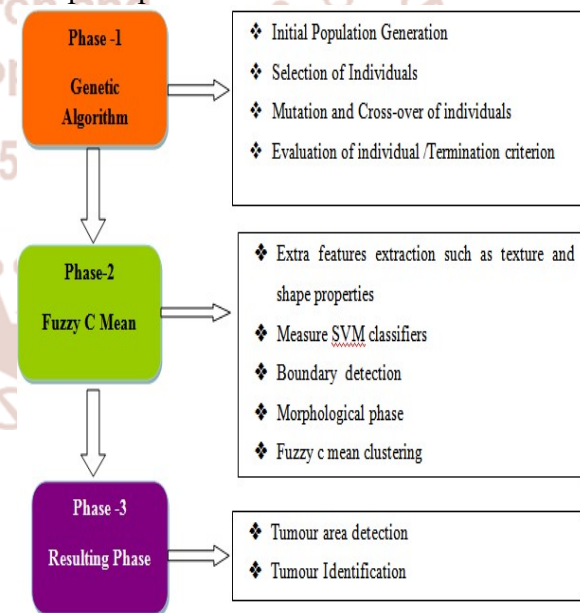


Figure 4.1 proposed HIMT model

The main key feature of proposed HIMT model is it assigns two or more than two clusters efficiently as compared to a K-means method. In k-mean method data point must exclusively belong to one cluster center and genetic algorithm is used as an optimization tool which helps to achieve results in less time.

In Genetic algorithm each of the chromosomes in the complete population of GA encodes a possible partition or clustering of the image set and the complete goodness of the chromosome is computed by using a fitness function. The main logic behind the integration are to reduce the total number of iterations or round done by variable initialization from the correct cluster centers to a correct Fuzzy C-means clustering method, it minimizes the total execution time and provides qualitative results.

**4.1 PROPOSED HINT ALGORITHM-**

HINT Algorithm for Tumour Detection from MRI Images-

**Input-** Set of MRI tumour Images

The input image F(i, j), where i, j are the pixel coordinates in the image.

**Output-** Tumor areas are detected for each input images, pixel k\*1, where k and l are pixel position of tumor

Step 1 - (Initialization) Select input image from image set (MRI).

Step 2 - Generates Initial Population (Genetic Algorithm)

Step 3 - Assign Membership degree using FCM

3.1 Assign Fitness

$$Jm(U, V) = \sum_{k=1}^m \sum_{i=1}^C (Uik)^m d_{2ik}$$

Step 4 - Noise removal by Median filter-

4.1 Select a noisy MRI image and separate each plane and each scalar component is treated independently.

4.2 Generate a zero based arrays near to image based on image mask size using pad array command.

4.3 Select a 3 \* 3 odd size image masks from an image set.

4.4 Then sort the pixel values within the mask in ascending order.

4.5 Under the mask, for each point, a single median component is determined.

Step 5 - Contrast enhancement-

5.1 Apply histogram technique (which maintains gray levels of the image)

Step 6 - Feature Extraction

6.1 Boundary or edge detection

6.1.1 The Intensity Gradient and Texture map are calculate

6.1.2 Extra features extraction such as texture and shape properties

6.1.3 Measure SVM classifiers

Step 7 - Final Phase

7.1 Apply Fussy C-Means clustering algorithm-include FCM (fuzzy c-means) and GFCM (geostatistical fuzzy c-means clustering) with marker watershed segmentation algorithm optimized C-means clustering method using SVM Techniques.

$$Q = \sum_{i=1}^C \sum_{j=1}^n umij |xj - oi|$$

Step 8 - Repeat steps 2 to 7 on each feedback

Step 9 - Generates final images.

**5. IMPLEMENTATION & RESULT ANALYSIS**

Simulations of existing and proposed method MATLAB tool are used. The confusion matrixes are used to evaluate and determines the performance of the proposed method. Here, two methods existing method (K- Means Clustering with GA [3]) and proposed method (HINT) are implemented. The proposed HINT method uses to combine the strategy of Fuzzy C-mean clustering, SVM classifier with Genetic algorithm. Following parameters were calculated for performance comparison-

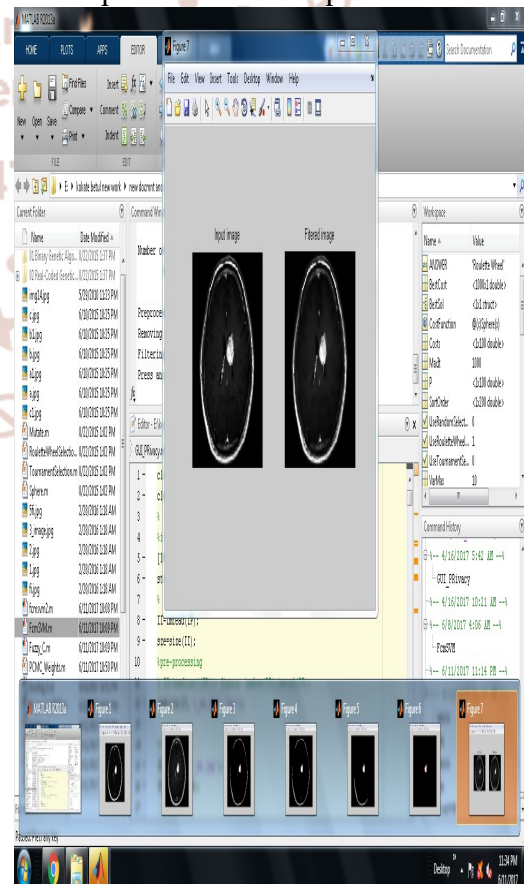
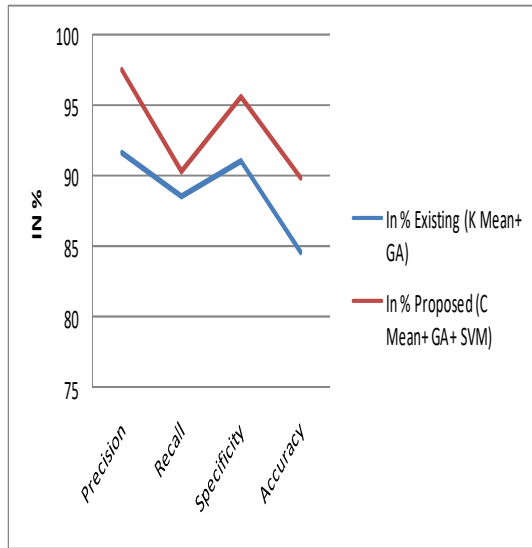


Figure 5.1 Simulation Snapshot

**5.1. PRECISION, ACCURACY, RECALL AND SPECIFICITY RESULTS-**

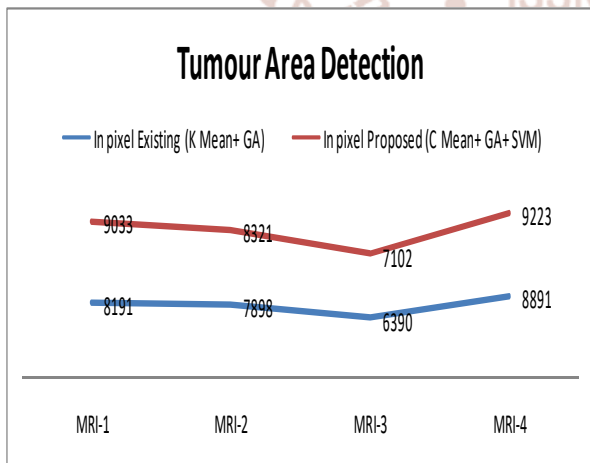
For existing (K Mean+ GA) and proposed HIMIT (C Mean+ GA+ SVM)-



Graph 5.1 Precision, Recall, Accuracy and Specificity Results- For Existing and Proposed

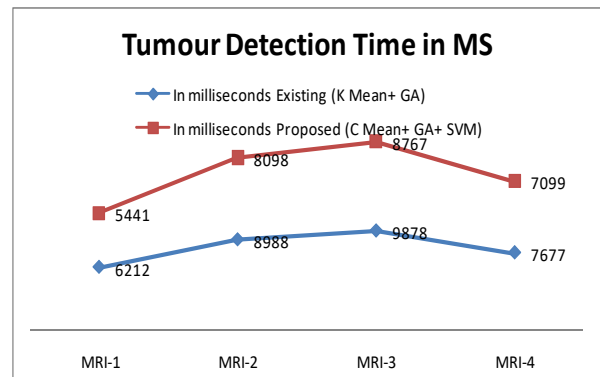
**5.2 TUMOUR DETECTION REGION-** Area of an image is calculated by knowing the vertical and a horizontal image resolution for an image.

Area of tumour=vertical resolution\* horizontal resolution\* Total number of image pixels in infected area



Graph 5.2 Tumour detection regions (Existing & Proposed)

**5.3 DETECTION TIME-** Searching time gives us how much time it takes for a segmentation method to generate the output.



Graph 5.3 Searching or detection Time

**5.4 RESULT COMPARISON AND ANALYSIS-**

Following comparison, parameters are calculated in between existing (K Mean+ GA)

Comparison parameters	Existing (K Mean+ GA)	Proposed (C Mean+ GA+ SVM)
Precision	Average	Better
Recall	Average	Better
Specificity	Average	Better
Accuracy	Average	Better
Area detection	Average	Better
Detection Time	Average	Better

and proposed HIMIT (C-Mean+ GA+ SVM).

Table 5.4- Result Comparisons

**Influences-** The above table 5.4 clearly shows proposed method performing outstanding over the existing image mining method in terms of various performance measurement parameters.

**6. CONCLUSIONS & FUTURE WORKS**

In this complete research work of image mining, we have presented a new hybrid model for image mining technique (called as HIMIT) for MRI medical images, to detect tumour more efficiently and accurately. The above results clearly show that, proposed HIMIT method performs outstanding over existing (K-Means+ GA) based method.

For future work, the set of (for different age group patients) brain medical images should be increases and use the real ones. By increasing the various number of image set, more training data set can be performed upgrade the system to the application for different size brain medical images slice. By doing this, the same area of the tumor but in the different slice of brain medical image can be determined. Also more no of feature can be extracted from images by using the 3-D slice for tumor detection.

## REFERENCES

1. G Rajesh Chandra, Dr. Kolasani Ramchand Rao, "tumour detection in brain using genetic algorithm", 7th International Conference on Communication, Computing and Virtualization 2016, Science Direct 449-457
2. M. G. Sumithra, B. Deepa, "Performance Analysis of Various Segmentation Techniques for Detection of Brain Abnormality", IEEE Proceeding of Region 10 Conference (TENCON) 2016, PP 2056-2061.
3. Anupurba Nandi, "Detection of human brain tumour using MRI image segmentation and morphological operators", 2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS), PP 55-61
4. R. preetha and G R Suresh "performance Analysis of Fuzzy C-Means Algorithm in Automated Detection of Brain Tumour" World Congress on Computing and Communication Technologies, Pages 30-33, 2014
5. Ramadass Sudhir "A Survey on Image Mining Techniques: Theory and Applications", Computer Engineering and Intelligent Systems, Pages 44-52, 2011
6. B. Jyothia, Y. MadhaveeLathab, P. G. Krishna Mohan ,V. S. K. Reddy," Integrated Multiple Features for Tumor Image Retrieval Using Classifier and Feedback Methods", International Conference on Computational Modeling and Security (CMS 2016), Science Direct, 2016, PP 141-148.
7. Müller H, Michoux N, Bandon D, and Geissbuhler A. A review of content-based image Retrieval systems in medical applications-clinical benefits and future directions. Medical Informatics. 1,73 (2004).
8. L.A. Khoo, P. Taylor, and R. M. Given-Wilson, "Computer-Aided Detection in the United Kingdom National Breast Screening Programme: Prospective Study," Radiology, vol. 237, pp. 444-449, 2005.
9. B. Jyothi, Y. MadhaveeLatha, P. G. Krishna Mohan," Multidimensional Feature Vector Space for an Effective Content-Based Medical Image Retrieval 5th IEEE International Advance Computing Conference (IACC-2015), BMS College of Engineering Bangalore, June 12 to 13, 2015.
10. Ashwani Kr. Yadav, R. Roy, Vaishali and Archer Praveen Kum," Survey on Content-based Image Retrieval and Texture Analysis with Applications" International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 7, No. "6 (2014), pp. 41-50.
11. Nitin Jain & Dr. S. S. Salankar," Color & Texture Feature Extraction for Content-Based Image Retrieval", IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE) e-ISSN: 2278-1676, p-ISSN: 2320-3331 PP 53-58.
12. Cs Pillai, "A Survey of Shape Descriptors for Digital Image Processing", IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555 Vol. 3, No.1, February 2013.
13. Ch. Srinivasa Rao, S. Srinivas Kumar, B.Chandra Mohan,"Content-based image retrieval using exact legend moments ad support vector machine", a the international journal of multimedia and its applications(IJMA), vol 2, no 2, may 2010.
14. Darshana Mistry," Survey of Relevance Feedback Methods in Content-Based Image Retrieval", Darshana Mistry / International Journal of Computer Science & Engineering Technology (IJCSET), Vol. 1 No. 2, pp 32-40, ISSN: 2229-3345.