# Performance Enhancement of Content Based Medical Image Retrieval for MRI Brain Images Based on Hybrid Approach

***G.Srilakshmi*[1]** *, Associate Professor, ECE, BRECW, Hyderabad, Telangana, India.*

***Dr.K.Rama Linga Reddy*[2]***, Head of the Department, ETM, GNITS, Hyderabad, Telangana, India.*

***Abstract—Content Based Image Retrieval (CBIR) systems retrieve images from the database which are similar to the query image. Instead of text based searching, CBIR efficiently retrieves images that are visually similar to query image. CBIR plays a major role in different domains including medical diagnosis, , industry inspection, geographical information satellite systems(GIS systems), biometrics, web based searching and historical research and so on. Content based image retrieval system is valuable in medical systems as it provides retrieval of the medical images which belongs to similar disease. Many diseases are curable, when they are diagnosed at the earlier stage. Diagnosis of disease depends on how efficiently the query image abnormality is accurately detected. The computer aided automated system such as content based medical image retrieval technique is used to retrieve query based images in the large database using combination of feature extraction and similarity matching methods. This paper aims to provide an efficient medical image retrieval in diagnosis of different brain diseases.The success of such a system largely depends upon the complexity, robustness , accuracy, and speed of the retrieval system.***

***Keywords—CBIR, Skull Stripping, GLCM, Wavelet Transform, PCA, GA, KNN.***

## I. INTRODUCTION

The field of medical imaging is gaining its importance with increase in the need of automated and efficient diagnosis in a short period of time. Computer and Information Technology are very much useful in medical image processing, medical analysis and classification. Accurate result can be produced only with the help of automated computer aided technique. CBIR can be useful for many diseases such as brain tumor, lung cancer, breast cancer, spine disorder problem etc. which is acquired through many modalities such as CT scan, MRI, mammogram etc. Medical images obtained by Magnetic Resonance (MR) imaging is used as a valuable tool in the clinical and surgical environment because of its characteristics like superior soft tissue differentiation, high spatial resolution and contrast. It does not use harmful ionizing radiation to patients [1].

Magnetic Resonance Images are examined by radiologists based on visual interpretation of the films to identify the presence of tumor abnormal tissue. The shortage of radiologists and the huge volumes of MRI to be analyzed make such readings labour intensive, cost expensive and often inaccurate. The sensitivity of the human eye in interpreting large numbers of images decreases with increasing number of cases, particularly when only a small number of slices are affected. Hence there is a need for automated systems for analysis, classification and retrieval of such medical images. The MRI may contain both normal slices and defective slices. The defective or abnormal slices are identified and separated from the normal slices and then these defective slices are further investigated for the detection of tumor tissues.

In this paper we concentrate on MRI brain images because brain is an important sense organ which controls and coordinate the whole homeostatic body function include heart beat , blood circulation etc. So the accurate retrieval of brain images is an important part of query based image retrieval from a large database[7]. Retrieval of similar images is based on similarity matching measures between features of query image with the database images features. The first step in this method is preprocessing because noise or misalignment will occur usually while capturing the images. After that feature extraction takes place such as content of image is represented using set of feature to avoid large input to the processing device, as all the content feature does not contain the useful information. The redundant features are reduced using some of feature selection method[2]. Next step is to compare the features of both query and stored images using distance measure and the similar images are retrieved with most similar image with lowest rank and so on.

This paper presents an hybrid approach for MRI image retrieval. Section II contains the methodology and details of extracted features, selection of features and distance measure of image features. Section III presents results and discussions. Section IV gives future scope in this field of medical image retrieval. Section V summarizes challenges in medical image retrieval. Section VI contains several concluding remarks.

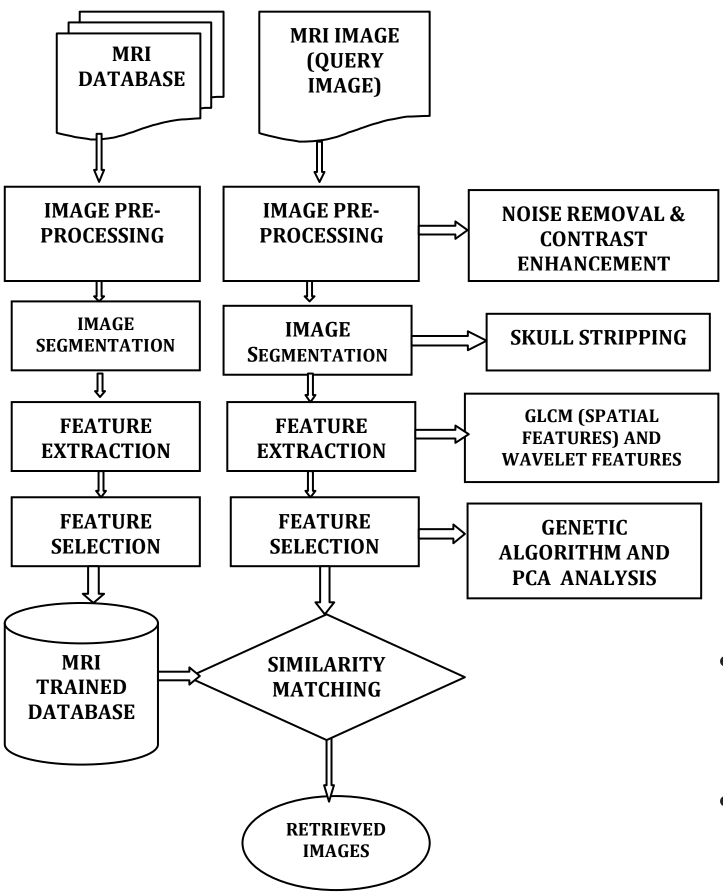


**Figure 1:** Proposed CBIR system for MRI images.

II. METHODOLOGY

The proposed methodology consists of a set of stages starting from collecting brain MRI images. The main steps are shown in Figure 1. This hybrid technique involves training and testing phase. Training phase consists of the following main steps : enhancement, skull stripping, segmentation, feature extraction using GLCM and DWT , storing in the database. Testing phase consists of enhancement, skull stripping, segmentation, feature extraction using GLCM and DWT. Next step is to compare the features of both query and stored images using distance measure and the similar images are retrieved with most similar image with lowest rank and so on. The images are processed through following stages:

- Acquisition of images
- Enhancement of MRI images
- Skull Stripping
- Feature Extraction
- Feature Selection
- Similarity Matching

*A. Acquisition of images*

Brain MRI images were collected and are converted into two dimensional matrices using MATLAB Software.

*B. Enhancement of MRI images*

In image preprocessing stage the noise is removed and contrast is enhanced. For removing high frequency noises, low pass filter is used. Most commonly weighted median filter (WMF) is used. Also for contrast enhancing, histogram stretching is chosen. The qualities of images are improved using enhancement technique. It is essential to improve the image information for human viewers, so that accurate segmentation outcomes are achieved. The methods given below are used for enhancement of brain MRI images.

- Contrast improvement- MRI images are RGB images, which are converted into gray scale images. These gray scale values are mapped into low and high intensity values using imadjust (MATLAB function).
- Mid-range Stretch- this is also an enhancement technique. In this method, the middle range MRI image intensity values are stretched. So it improves the quality of brain MRI images. In this technique, gray scale image pixels are mapped between 0 and 1

 

value by dividing 255 intensity values as shown below:

$$X_{ij}= I_i/255 \quad (1)$$

Here i for row index of brain image matrix and j for column.To compute the function f(x) on the X matrix obtained from equation (1). The function f(x) is defined as follows.

$$f(x_{ij}) = \begin{cases} 0.5 * x_{ij} & x_{ij} < 0.1 \\ 0.1 + 1.5 * (x_{ij} - 0.1) & 0.1 \leq x_{ij} \leq 0.88 \\ 1 + 0.5 * (x_{ij} - 1) & x_{ij} > 0.88 \end{cases} \quad (2)$$

Subsequently applying the above function f ($x_{ij}$) the gray-scale images are converted to indexed images. The output images obtained after applying all the operations are done to improve the quality of images.

*C. Skull Stripping*

In Image segmentation stage, undesired tissues such as skull, nose and eyes are deleted. Skull stripping is a significant step. The steps involved in skull stripping are given below.Double thresholding- it is a segmentation technique. This technique, convert the image into binary form, that is gray scale image to binary image. This technique generate the mask by setting each pixel in the range of 0.1*255-.88*255 to 1 means white and remaining pixels to 0 means black. Non brain tissues pixels were discarded in MRI image. Here two thresholds upper and lower are considered so it is known as double thresholding technique .

- Erosion- in this stage unwanted pixels are removed from MRI image after thresholding. Thus the skull portions are removed. Here disk of radius 3 was taken as a structuring element for removing all unwanted pixels which are contributing to the brain MRI images.
- Region filling- this method is used to fill the holes in the images. After the erosion, eroded images are filled using region filling algorithm. Here the associated background pixels are converted into foreground pixels so that the holes present in the eroded images are removed in brain MRI image.

*D. Feature Extraction*

The purpose of feature extraction is to reduce the original data set by measuring certain properties, or features, that distinguish one input pattern from another. The extracted features provide the characteristics of the input type to the classifier by considering the description of the relevant properties of the image into a feature space. The texture analysis to the T1, T2 and T2 FLAIR images to describe quantitatively the brightness and texture of the images is used. Texture analysis covers a wide range of techniques based on first- and second order image texture parameters.

*(i)Grey Level Co-Occurrence Matrix (GLCM)*

The GLCM was introduced by Haralick et al. [3]. It is a second order statistical method which is reported to be able to characterize textures as an overall or average spatial relationship between grey tones in an image [4]. The second order probabilities were sufficient for human discrimination of texture. In general, GLCM could be computed as follows. First, an original texture image I is re-quantized into an image G with reduced number of grey level, $N_g$. A typical value of $N_g$ is 16 or 32. Then, GLCM is computed from G by scanning the intensity of each pixel and its neighbor, defined by displacement d and angle ø. A displacement, d could take a value of 1,2,3,...n whereas an angle, ø is limited 0°, 45°, 90° and 135°.

The GLCM P(i,j,d,ø) is a second order joint probability density function P of grey level pairs in the image for each element in co-occurrence matrix by dividing each element with $N_g$. Finally, scalar secondary features are extracted from this co-occurrence matrix. All these features were employed as inputs to feature selection.

$$\text{Energy}: \sum_{i,j} P(i,j)^2 \quad (3)$$

$$\text{Entropy}: -\sum_{i,j} P(i,j)\log P(i,j) \quad (4)$$

$$\text{Homogeneity}: \sum_{i,j} \frac{1}{1+(i-j)^2} P(i,j) \quad (5)$$

$$\text{Inertia}: \sum_{i,j} (i-j)^2 P(i,j) \quad (6)$$

$$\text{Correlation}: -\sum_{i,j} \frac{(i-\mu)(j-\mu)}{\sigma^2} P(i,j) \quad (7)$$

$$\text{Variance}: \sum_{i,j} (i-\mu)^2 P(i,j) \quad (8)$$

where $\mu$ is mean given by $\sum_{i,j} P(i,j)$

*(ii)Wavelet Features*

Also, DWT reduces redundancies but still these features are numerous and redundant. Usually after Wavelet transformation, a feature selection algorithm, PCA, is used to reduce feature numbers, so at last features with lowest correlation remain at last. Histogram is another tool to extract image features. Here histogram diagram is divided to 8 intervals. The values are normalized and then entered in matrix as features.

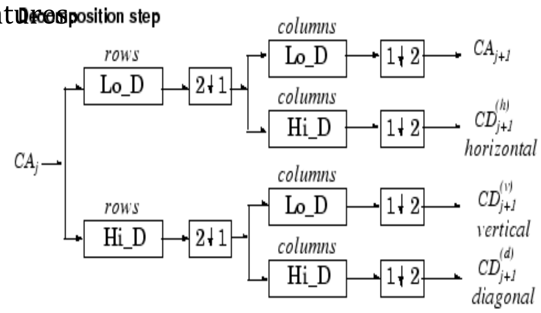


Figure 2- Two level DWT of original image.

The figure-2 shows the two level decomposition of Haar DWT with approximate and detail co-efficients.

*E. Feature selection*

In this paper Genetic algorithm (GA) is used for feature reduction. GA is realized as computer program that has been very successful in many optimization areas. GA emulates chromosomes, reproduction and selection. It creates a population of individuals. The computer program evaluates each individual and gives it a degree of fitness. Individuals are given opportunities to reproduce proportionally to their fitness. The fitness function is crucial factor in the success of the GA [3]. In this paper, the proposed algorithm in is used for feature reduction using GA. The applied GA fitness function is based on the principle of Max-Relevance and Min-Redundancy [5]. The proposed algorithm is based on mutual information, so that the feature selection is to find a feature set S with m features {x,}, which jointly have the largest dependency on the target class c.

*F. Distance Measure using KNN (K_Nearest _Neighbors)*

K Nearest Neighbors is a supervised machine learning method, which stores all available cases and classify new cases based on a similarity measure. The algorithm is based on classification using a majority vote of its neighbors, with the case being assigned to the class most common amongst its K nearest neighbors measured by a Euclidean Distance.

*G.Performance measures*

Classification sensitivity, specificity and accuracy are calculated using below formulae:

- True Positive (TP): Abnormal brain correctly identified as abnormal.
- True Negative (TN): Normal brain correctly identified as normal.
- False Positive (FP): Normal brain incorrectly identified as abnormal.
- False Negative (FN): Abnormal brain incorrectly identified as normal.

1) Sensitivity = TP/ (TP+FN) *100%
2) Specificity = TN/ (TN+FP) * 100%
3) Accuracy = (TP+ TN)/ (TP+ TN+FP+FN)* 100 %

All these three parameters are used to check the classifier performance.

*4) Retrieval Efficiency*

The retrieval efficiency, namely recall and precision are calculated.

$$precision = \frac{No.\ of\ relevant\ images\ retrieved}{Total\ No.\ of\ images\ retrieved}$$

$$recall = \frac{No.\ of\ relevant\ images\ retrieved}{Total\ No.\ of\ relevant\ images\ in\ the\ database}$$

The table-1 below is the contingency table of the classifier.

**Table 1. Contingency table**

| Actual Group | Predicted Group | |
|---|---|---|
| | **Normal** | **Abnormal** |
| Normal | TN | FP |
| Abnormal | FN | TP |

**III. RESULTS AND DISCUSSIONS**

The input database in this consists of three sets of abnormal MRI data. The first set contains synthetic brain MR images obtained from BrainWeb, Simulated Brain Database at the McConnell Brain Imaging Centre of the Montreal Neurological Institute (MNI), McGill University (http://www.bic.mni.mcgill.ca/brainweb). No image registration was required in the case of synthetic images, but image registration using Matlab functions is applied on clinical dataset as a preprocessing step. Implementation of step 2 in proposed algorithm is done with a 4-level wavelet decomposition of input signals. Mathworks Matlab 13.0 (R2013a) implementation on a PC with Pentium Dual CPU of 2.0GHz and 2GB RAM running Microsoft Windows 7 was executed for the complete system evaluation.

 

**Table 2 :** Classification Accuracy comparison with other Methods

| Reference No. | Approach | Classification Accuracy (%) |
|---|---|---|
| [5] | GLCM+ PCA + SVM | 95 |
| [6] | DWT + PCA + ANN | 98.33 |
| Proposed method | DWT+PCA+GA+KNN | 98.43 |

The table-2 gives the comparison of classification accuracy of other two approaches with the proposed method. The proposed method gives classification accuracy of 98.43 %.

## IV. FUTURE WORKS

This proposed method has been tested with the database of MRI brain images. Today we have ample of fast speed processors and cheap storage but as amount of images increases complexity also increases. We can use indexing in brain images, and other mathematical notations like Features Energy, Entropy, Contrast, Inverse Difference Moment, Variance, Sum Average, Sum Entropy, Sum Variance, Difference Variance, difference Entropy or combination of any of them to deal with this increasing complexity. Image localization can also be added in preprocessing part so that system will focus on the consistent particular region of interest only and give more refined results.

## V. CHALLENGES IN MEDICAL IMAGE RETRIEVAL

Some of the major challenges in the area of medical image retrieval are outlined as follows:

1. Extraction of robust and precise visual features from medical images is a difficult problem.
2. The use of CBIR in medical diagnostics is important though it is difficult to realize.
3. To be used as a diagnostic tool, the CBIR systems need to prove their performance to be accepted by the clinicians.
4. In medical application domain many systems have been proposed where database consists of images of various anatomical regions with variety of image modalities. The standard databases are useful as a benchmark to test various approaches in a general image retrieval framework; however these approaches are not useful for diagnostics support systems where high precision is required. Useful semantics for medical image retrieval needs to be established[2].

## VI. CONCLUSION

In this proposed system , hybrid methodology of combining genetic algorithm and feature selection after rigorous preprocessing and feature extraction techniques give accurate result for retrieving all similar brain diseased MRI images.

## REFERENCES

[1] D. Jude Hemanth, C. Kezi Selva Vijila, J. Anitha , "Artifical intelligent based Brain Pathology Identification techniques in Magnetic resonance Images." International journal of reviews in computing. 2009.

[2] Y. Fanid Fathabad, M.A. Balafar, "Content Based Image Retrieval in Medical", International Journal on Technical and Physical Problems of Engineering (IJTPE), Issue 12, Vol. 4, No. 3, pp. 177-182, September 2012.

[3] H. Muller, et al., "A Review of Content Based Image Retrieval Systems in Medical Applications Clinical Benefits and Future Directions", International Journal of Medical Informatics, Vol. 73, No. 1, pp. 1-4, 2004.

[4] Pratik Vinayak Oak and R. S. Kamathe, "Contrast Enhancement of Brain MRI Images using Histogram Based Techniques ", International Journal of Innovative Research in Electrical, Electronics, Instrumentation, and Control Engineering, Vol.1, Issue 3, June 2013.

[5] Mina Rafi Nazari and Emad Fatemizadeh, 'A CBIR

 

System for Human Brain Magnetic Resonance Image', International Journal of Computer Applications (0975 – 8887) Volume 7– No.14, October 2010 Indexing.

[6] El- Shayed A El Dahshan, Tamel Hosny, Abdel-badech M. Salem, "Hybrid intelligent techniques for MRI brain images classification", Elsevier Journal of Digital Signal Processing, Vol. 20 , No 2 ,pp 433-441,2010.

[7] G.Sharmila, D.Saraswathi, R.Ramkumar "A SURVEY ON CONTENT BASED MEDICAL IMAGE RETRIEVAL FOR MRI BRAIN IMAGES", International Journal of Research in Engineering and Technology Vol. 03 Issue 07, May-2014.

BIOGRAPHIES:

G.Srilakshmi received her B.Tech degree from GNITS, Hyderabad in the year 2001 and M.Tech degree from JNTUH in the year 2005. Her teaching experience is 10 years. Presently working as an Associate Professor at Bhoj Reddy Engineering College for Women, Hyderabad. Her research interests include Biomedical Image Processing and Signal Processing.

Dr.K.Rama Linga Reddy completed his B.E. from Vasavi Engineering College in 1989 , M.Tech from SVU Engineering College in 1991 and Ph.D from JNTUH, Hyderabad. He has 23 years teaching experience (15 years in GNITS & 8 years in CBIT).He has 30 research papers in his credit and has participated in 40 T.V programs (Panel Discussions related to quality of engineering education and EAMCET counseling) in different T.V news channels. He worked as advisor for 15 engineering colleges in NBA accreditation process. He visited South East Asia Countries. Currently working as Head of the Department in Electronics and Telematics (ETM) department in G.Narayanamma Institute of Technology and Science (GNITS), Hyderabad.

 