MRI Brain Image Classification and Detection Using Distance Classifier Method in Image Processing

P.B.Nikam
M.E. Student, S.S.V.P.S's B.S.D.C.O.E,
Dhule, Maharashtra.

V.D.Shinde
Asst. Professor S.S.V.P.S's B.S.D.C.O.E,
Dhule, Maharashtra.

Abstract: - Human brain is a very complex structure and Brain image analysis is a very intricate task. Detection and segmentation of tumors from brain is very difficult due to variance and complexity of tumors and dense brain tissues. Manual segmentation of these abnormal tissues may result in misdiagnosis due to human errors. For this automatic medical image segmentation and classification is a key step in computer aided imaging to get accuracy and correct estimation of medical images. This paper presents a system for automatic classification of healthy or affected person using Region growing segmentation by watershed algorithm, Euclidean distance classifier for fast computation, accompanied with preprocessing and post processing method apply on database consisting both normal and tumorous samples of MR brain images in jpg format, this method is implemented using MATLAB version 10.The results ensures that the method is efficient, and satisfying for quick detection whether person is healthy or unhealthy.

Key words: Euclidean distance classifier, GLCM, MRI, watershed segmentation.

INTRODUCTION

The human brain is possibly one of the most complex systems in the universe. Nowadays various technologies are used for imaging and recording brain activity for diagnosis of various abnormalities e.g., electroencephalograms (EEG), magneto encephalograms(MEG),X-RAY ,magnetic resonance imaging (MRI),functional magnetic resonance imaging (fMRI), computer tomography (CT), positron emission tomography (PET), etc. Those brain imaging tools allow researchers to gain understanding of the complex inner mechanisms of the brain. The various imaging methods such as x-ray, magnetic resonance, computed tomography, ultrasound, fusion of modalities etc are very important in today's medical imaging[2][3] .Out of which MRI is widely used for brain imaging. Due its advantages over other imaging techniques it provides one visualization of the intracranial circulation and extra cranial carotid vessels, etc. It gives soft tissue contrast and noninvasiveness. MRI also provides unparallel view inside the human body. It gives extraordinary level of detail compared with other modalities, no ionizing radiation, and better spatial resolution over horizontal, frontal or sagittal planes are others advantages MRI can be done number of times a day. To properly analyze this data various image processing tools are required like enhancement, denoising, edge detection, morphological operations, segmentation, feature extraction, classification etc. These all tools up to classification creates a medical database used for learning and diagnosis purpose which is useful for different modalities that are taken under variable conditions with variable accuracy[5]. The MRI is most commonly used in diagnosis and prognosis of various brain abnormalities as brain tumor, Alzheimer's disease, Szerophrenia, ischemia, sclerosis, epilepsy etc .but the project deals with diagnosis of person with or without tumor. Brain tumor is an abnormal intracranial growth caused by cells reproducing themselves in an uncontrolled manner which accounted for 1.4 percent of all cancers, 2.4 percent of all cancer deaths, and 20 to 25 percent of pediatric cancers. It was estimated that there are approximately 13,000 deaths per year as a result of brain tumors. Brain tumors are the leading cause cancer death in children under the age of 20. They are the second leading cause of cancer death among 20-29 year old males [1].

I. DATABASE

Database consists of real data images of MRI brain scan of human being with tumor and without tumor. The MRI brain images are taken of person above 18 years age group as there are no significant changes observed in image pattern after 18 years of age. Database consists of images of both male and female collected from various sources as from doctors, MRI centers in DICOM image format and from internet in JPG, JPEG, and other image format of variable sizes up to 400kb.

II. METHODOLOGY

This system had two main stages, first is pre-processing of MRI images and then other post processing operations like segmentation, edge detection, morphological operations feature selection and extraction, classification etc. Steps of algorithm are as following:[14][15]

Stage 1: Pre-processing: Pre-processing includes operations like noise removal, input image into gray scale image it includes: steps from i to iii.

i.) Gray scale imaging: Grayscale is a range of shades of gray without any apparent color. The darkest shade in gray scale is black which is due to total absence of transmitted or reflected light. White is the lightest shade which is due to the total transmission or reflection of light at all visible wavelengths. MRI is magnetic resonance imaging acquired by MRI machine of the part of the body which is under test or desired here we are interested in brain. In analog practice, gray scale imaging is also called as black and white, but in digital imaging. Images can be either true black or white it is also known as halftone, The effect of gray shading in a halftone image is obtained by depiction of the image as a grid of black dots on a white background (or vice versa), the sizes of the individual dots/pixels determine the apparent lightness or darkness representing gray shade in their locality.

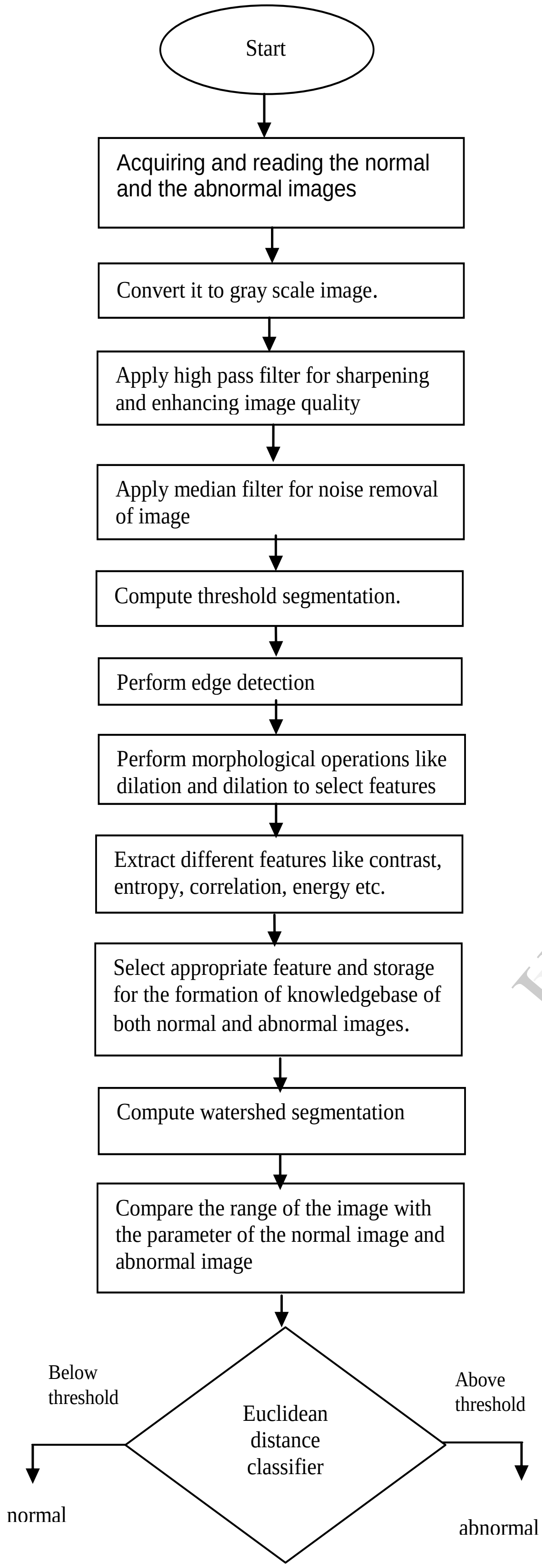


Fig 1 Flowchart for the system

ii.) High pass filter: Human perception is highly sensitive to edges and fine details of an image, and since they are composed primarily by high frequency components, the visual quality of an image can be enormously degraded if the high frequencies are attenuated or completed removed. In contrast, enhancing the high-frequency components of an image leads to an improvement in the visual quality. Image sharpening refers to any enhancement technique that highlights edges and fine details in an image. Image sharpening consists of adding to the original image a signal that is proportional to a high-pass filtered version of the original image. Often referred to an unsharp masking on a one-dimensional signal. The original image is first filtered by a high-pass filter that extracts the high-frequency components, and then a scaled version of the high-pass filter output is added to the original image, thus producing a sharpened image of the original. The kernel of the high pass filter is designed to increase the brightness of the center pixel relative to neighboring pixels. The kernel array usually contains a single positive value at its center, which is completely surrounded by negative values. [9][10]

iii.) Median Filter: In MRI brain images, there are chances of having presence of salt and pepper noise, it is desirable to perform noise reduction on an image. The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a very important step to improve the results of later processing (for example, edge detection on an image). Median filtering is a nonlinear operation often used in image processing to reduce "salt and pepper" noise. A median filter is more effective than convolution when the goal is to simultaneously reduce noise and preserve edges. The median filter is a sliding-window spatial filter. It replaces the value of the center pixel with the median of the intensity values in the neighborhood of that pixel. For every pixel, a 3x3 neighborhood with the pixel as center is considered. In median filtering, the value of the pixel is replaced by the median of the pixel values in the 3x3 neighborhood. If the window has an odd number of entries, then the median, it is just the middle value after all the entries in the window are sorted numerically. For an even number of entries, there is more than one possible median. This filter thus enhances the quality of the MRI image. [9]

Stage 2: Post processing: Pre-processing techniques enhances the image and make it suitable for further operations which are very important for performing to do the task of diagnosis by post processing techniques.

iv.)Threshold segmentation: Image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain visual characteristics. 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. Thresholding is the simplest way to perform segmentation, and it is used extensively in many image processing applications. Thresholding is based on the notion that regions corresponding to different regions can be classified by using a range function applied to the intensity values of image pixels. The assumption is that different regions in an image will have a distinct frequency distribution and can be discriminated on the basis of the mean and standard deviation of each distribution. Several popular methods are used in biomedical applications are global thresholding, maximum entropy method, Otsu's method (maximum variance), k-means clustering can also be used. In computer vision, Segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify or convert the representation of an image into something that is more meaningful, detail and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.)[6][9]

v.) Edge detection: - Edge detection is the most common approach for detecting meaningful discontinuities in gray level. Edge is a set of connected pixels that lie on a boundary between the two regions. Edges are more closely modeled as having ramp like profile. It is observed that the edges pixels are more reduced for tumor image as compared to a non-tumor image. This reduction of edge count for tumor images is because the brain substance has been pushed aside and compressed by the growth of the tumor within a confined space i.e. the intracranial cavity. Such symptoms occur due to the local effects of compression or invasion on cerebral tissue .it is reported by Neurosurgeons that systemic malignancy can metastasize to any location in the brain but most commonly (80%) affects the cerebral hemispheres, this accounts the tendency of metastases located at the junction of the gray and white matter thus occurrence of cerebral compression reduce the edge counts of the cerebral hemisphere. From this understanding, we use the canny edge detection method to detect image edges (*IE*). The Canny operator works in a multi-stage process. First the image is smoothed by Gaussian convolution. Then a simple 2-D first derivative operator is applied to the smoothed image to highlight regions of the image with high first spatial derivatives. Edges give rise to ridges in the gradient magnitude image.[7][9].

vi.) Morphological operations: Mathematical morphology is a tool for extracting image components that are useful in the representation and description of shape such as boundaries, skeletons, and complex hull. Morphological operations depend only on the relative ordering of pixel value rather than on their numerical values, and therefore are particularly appropriate for the processing of binary images. There are two fundamental morphological operations Dilation and Erosion, in terms of an image with a translated shape (structuring element). Morphological techniques probe an image with a small shape or template called a structuring element. a.) Dilation thickens objects in a binary image the specific extent of this element is controlled by structuring element. b.) Erosion: shrinks object in binary image. The strel function constructs structuring element with a variety of shapes and sizes. Its basic syntax is:

Se= strel (shape, parameter).

For this system development the strel functions used are

i. disk: creates a disk shaped structuring element with radius R (additional parameters may be specified for the disk, and

ii. Square: creates a square structuring element whose width is W pixels. W must a nonnegative integer scalar.

vii.) Feature extraction: Features gives the characteristics of the objects of interest, are representative of the maximum related information that the image has to offer for a complete depiction of a tumor. Feature extraction techniques analyze images to extract the most prominent features that are representative of the various classes of objects. Features are used as inputs to classifiers which assign them to the class that they represent. Feature extraction enable to reduce the original data by measuring certain properties of images which have relevant data, or features, that distinguish one pattern from another pattern. There different types of features like shape based, color based, and texture based etc, texture is one of the important characteristics used in identifying objects or regions of interest in an image. Some texture features used in this system are as given below, these features like contrast, energy, entropy, homogeneity; correlation etc. These parameters can be calculated by using graycoprops command, which also being an inbuilt function in MATLAB it calculates the statistics specified in properties from the GLCM. GLCM is an m-by-n-by-p array of valid gray-level co-occurrence matrices. If GLCM is an array of GLCMs, stats is an array of statistics for each GLCM. Graycoprops normalizes the gray-level co-occurrence matrix (GLCM) so that the sum of its elements is equal to 1. Each element (r, c) in the normalized GLCM is the joint probability occurrence of pixel pairs with a defined spatial relationship having gray level values r and c in the image. Graycoprops uses the normalized GLCM to calculate properties. [4][8][19]

Viii.) Feature selection: Feature selection is a process of selecting appropriate features from all extracted features as theoretically more features are preferred as it gives more selective power. But this is not practical as with an inadequate number of training data these can't produce satisfying results, that's why excessive features are omitted. Excessive features also slow the training process. Excessive features may also cause the classifier to over-fit the training data as inappropriate features may mislead the classifier. Feature selection approach does not attempt to generate new features instead it tries to select $d < D$ features by getting rid of inappropriate and unneeded features and maintaining the appropriate ones. In this project 5 features are selected as below: [4][8][19]

a. Contrast: Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. Range = [0 (size (CILCMJ)-l)"2) Contrast is 0 for a constant image.

$$\sum_{,} |-|_{p(i,j)}$$

b. Correlation: Returns a measure of how correlated a pixel is to its neighbor over the whole image. Range = [-1 1] Correlation is1or-1foraPerfectly positively or negatively correlated image. Correlation is NaN for a constant image.

ISSN: 2278-0181
Vol. 2 Issue 6, June - 2013

$$\text{Cor} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{ijp(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

c. Energy: Returns the sum of squared elements in the GLCM Range= [0 1] Energy is 1 for a constant image.

$$E = \sum_{I=0}^{G-1} p[(i)]^2$$

d. Homogeneity: Returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Range=[0 1] Homogeneity is 1 for a diagonal GLCM.

$$H = \sum_i \sum_j [M(i,j)]^2$$

e. Entropy: Measures the randomness of a gray-level distribution. The Entropy is expected to be high if the gray levels are distributed randomly throughout the image.

$$E = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} p(i,j) log_2 \, [p(i)]$$

ix.) Watershed segmentation: The watershed transformation considers the gradient magnitude of an image as a topographic surface. Pixels having the highest gradient magnitude intensities (GMIs) correspond to watershed lines, which represent the region boundaries. Water placed on any pixel enclosed by a common watershed line flows downhill to a common local intensity minimum (LIM). Pixels draining to a common minimum form a catch basin of the drops of water, which represents a segment. In image processing, different watershed lines may be computed. In graphs, some may be defined on the nodes, on the edges, or hybrid lines on both nodes and edges. Watersheds may also be defined in the continuous domain There are also many different algorithms to compute watersheds.[9]

x.) Euclidean distance classifier: The most familiar point to distance measure is Euclidean distance. The classifier based on this distance measure is direct and simple. The mean class values are used as class centers to calculate pixel-center distances for use by the Euclidean distance rule. For major level classification of a homogeneous area this scheme is better. Its advantageous nature comes from the minimum time it takes to classify Distance Measures are used to group or cluster brightness values together. Euclidean distance between points in space is a common way to calculate closeness Euclidean metric is the "ordinary" distance between two points that one would measure with a ruler, and is given by the Pythagorean formula.

$$d = |X - Y| = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$$

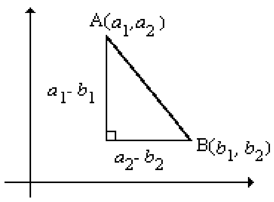


Fig 2. Euclidean distance classifier based on Pythagoras theorem

$$|AB| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$$

Euclidean classifier takes very lesser time when compared to other classifiers; still the accuracy attained with this method is good. From the results it is clear that this classifier is very quick as the classification of a data sample requires the less evaluation of the decision function for each class under consideration. For this size of the data set is immaterial to the process since each data point is classified independently.[15][16]

III. RESULTS AND DISCUSSIONS

Results for this project work are based on database of 25 images which contains 5 Normal or Healthy brain scan and 20 images of abnormal or scan of person affected from tumor. For this purpose real time patient data is taken for study. As tumor in MRI image have intensity more than that of its background so it become very easy locate it and extract it from a MRI image. The main purpose of the developed system is to detect whether person is healthy or unhealthy so to verify for the results and testing algorithm the different parameters are calculated. Following figures shows results of various operations.

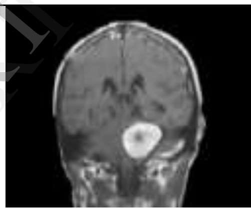
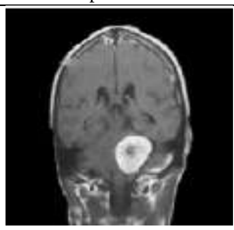
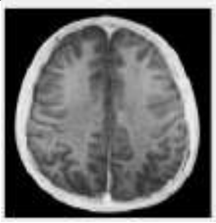
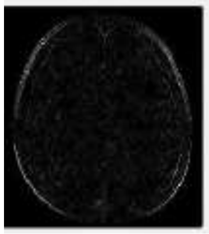
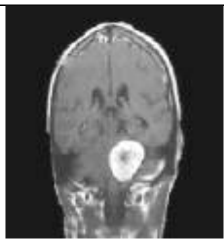
Anormal image	Normal image
 Fig 4a.Input image of tumor affected person	 Fig 4b. Input image of healthy person
 Fig 5a. rgb to grayscale image	 Fig 5b. rgb to grayscale image
 Fig 6a. HPF output of fig.5 image	 Fig 6b. HPF output of fig.5 image
 Fig 7a. Enhanced image of fig.6	 Fig 7b. Enhanced image of fig.6

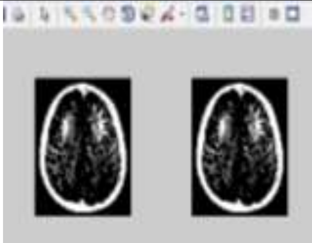
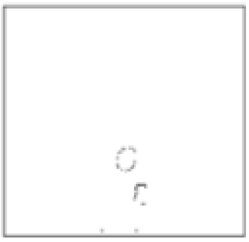
 Fig. 8a. Edge detection image of i/p image	 Fig. 8b. Edge detection image of i/p image
 Fig. 9a. Morphological operation	 Fig. 9b. Morphological operation
Feature extraction and selection for abnormal image Mean =12.0924 Std = 30.8336 Contrast= 428.9425 Correlation = 0.8954 Energy = 0.8427 Homogeneity = 0.9284 Entropy = 0.2962	Feature extraction and selection for normal image Mean = 7.1068 Std = 18.3163 Contrast = 332.1653 Correlation = 0.7718 Energy =0.8430 Homogeneity = 0.9329 Entropy =0.3264
 Fig. 10a. Watershed segmented image of i/p image	 Fig. 10b. Watershed segmented image of i/p image

Fig.11a. final result of image recognition of i/p image Fig. 11 b. final result of image recognition of i/p image

IV. FUTURE WORK

In future the system has to be checked for large database (above 100) MRI images to check the performance of the system whether it gives results with same speed as it takes very less time to give accurate results for this program now, also want to advanced this program to detect tumor growth and penetration by modifying this program.

V. CONCLUSION

This system was developed to diagnose that the image is of healthy person or unhealthy person using a simple but quick unsupervised classifier Euclidean distance classifier. From the results it is seen that it takes very less evaluation time few seconds to recognize whether image is normal or abnormal. This program is quick and gives results better and quick as compared to other supervised classifiers it is beneficial where large databases have to compute rapidly.

REFERENCES

[1] Ricci, P. E. and Dungun. D. H. (2001). Imaging of low- and intermediate-grade gliomas. *Seminars in Radiation Oncology, Vol 11, No 2 (April), 2001: pp 10i-112,* 1 1(2):103 - 114

[2] Lorena Tonarelli," Magnetic Resonance Imaging of Brain Tumor," Enterprises for Continuing Education Inc. (ECEI), PO Box 300, Brighton for www.*CEwebsource.com*

[3] N. Leeds and E. Jackson, "Current imaging techniques for the evaluation of brain neoplasm's," *Current Sci.*, vol. 6, pp. 254–261, 1994.

[4] M. Sasikalal and N. Kumaravel," Comparison of Feature Selection Techniques for Detection of Malignant Tumor in Brain Images," IEEE Indicon 2005 Conference, Chennai, India, I I - 1 3 Dec. 2005

[5] Phooi-Yee LAU and Shinji OZAWA, "A Simple Method for Detecting Tumor in T2-Weighted MRI Brain Images,": An Image-Based Analysis," Department of Information and Computer Science, Keio University, Yokohama-shi, Japan

[6] Matthew C.Clark, Lawrence O.Hall ," *Automatic tumor segmentation using Knowledge based techniques*," IEEE transactions on medical imaging, vol 17, No 2, April 1998

[7] Kulbir Singh 2006,"Edge based Region growing,"- dissertation report from Thapar University Patiala, India 9-21, 30-35.

[8] Monika Arrora June 2009,"Detection of abnormalities in MRI images using texture analysis," - dissertation report from Thapar University Patiala, India 45-65.

[9] R.C.Gonzalez, R.E.woods & S.L.Eddins , "Digital Image processeing using matlab" second edition.100-105,334-374,378-422

[10] Rajesh C. Patil, Dr. A. S. Bhalchandra," Brain Tumor Extraction from MRI Images Using MATLAB," International Journal of Electronics, Communication & Soft Computing Science and Engineering ISSN: 2277-9477, Volume 2, Issue 1

[11] W. Gonzalez, "Digital Image Processing", 2nd ed. Prentice Hall, Year of Publication 2008, Page no 378.

[12] Erlend Hodneland July 2003," Segmentation of Digital Images," – report University of Bergen. 8-17

[13] Lihong Zheng and Xiangjian He," Classification Techniques in Pattern Recognition," Conference proceedings ISBN 80-903100-8-7 WSCG'2005, January 31-February 4, 2005 Plzen, Czech Republic.

[14] Alessandro Bevilacqua," Pattern recognition classification and clustering,"-report for ACAT, Advanced School - Bologna, May 25, 2012

[15] Giovanni Luca Masala," Pattern Recognition Techniques Applied to Biomedical Patterns," International Journal of Biological and Life Sciences 1:1 2005

[16] Subash Chelliah,"A Study of Euclidean classifier," Paper for Remote Sensing Division ,Birla Institute of Scientific Research Jaipur, India

[17] V.P.Gladis Pushparathi, S.Palani, "*A novel approach for feature extraction and Selection on MRI images for brain tumor classification* ",Proc, CCSEA, SEA,CLOUD, DKMP, CS & ITCSC

[18] Smitha P, Shaji L. *"A review of medical image classification techniques"* (ICVCI) 2011

[19] Robi Polikar ," Pattern Recognition, "Wiley Encyclopedia of Biomedical Engineering, Rowan University Glassboro, New Jersey

[20] Thesis on brain image segmentation using graph cuts by department of Signal and Systems Chalmer university of technology, report no.ex072 / 2010

[21] Multimodality multiparameter medical image analysis 5th international conference(IJIST)vol.2;no4 July 2012

[22] Jiawei Han and Michelene Kamber book "data mining techniques, 2nd edition "pg(286-300)