Multiple Ranks Weighting Score for Microscopic Image Retrieval System

P. Suresh
M.Kumarasamy College of Engineering, Karur, India
Email: psureshengineer@gmail.com

L. Malliga
M.Kumarasamy College of Engineering, Karur, India
Email: malligamkce@mkce.ac.in

M. Vijay
M.Kumarasamy College of Engineering, Karur, India
Email: vijay.vicki@gmail.com

Abstract— Content based medical images have become a major necessity with the growing retrieving Advancements. CBIR access to medical images for supporting clinical decision making has been proposed that would be ease to manage large number of image in the database system. [4] In real time case only few systems has been developed and used in clinical environment. Content-Based Image Retrieval refers to image retrieval system that is based on visual properties of image objects other than textual annotation. Query image features compare with the database image features which is not exactly matching so image feature can be compare with the two tier approach in the database image in order to improve the accuracy of the retrieval system. Every day, large volume of different types of medical images such as MRI, CT images ultrasound, x-ray, radiology, etc are produced in different medical centre's .microscopic image classification and discrimination (sub-type) [12] is the most difficult problem in medical image retrieval system. In this paper, the survey provides the suitable algorithm for retrieval and classification of medical image to improve the overall accuracy of the MIMS.

Index Terms— Content Based Image Retrieval (CBIR), Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Medical Image Management System (MIMS)

I. INTRODUCTION

Medical images are composed into different minor structure so feature extraction and classification should be accurate [3] and easy to retrieve from the clinical database. Fast retrieval of image from large database is the most important factor that should be addressed. Basic characteristics of CBIR are high retrieval efficiency and less computational complexity. In conventional databases, image retrieved based on keyword because images are text-annotated so it has the following limitations.

- Keyword classification is very subjective approach
- There is no fixed set of key word which describes the content of the images

In order to overcome the manual annotation of the image, Content based image retrieval has introduced which describes the image based on text, colour and size or shape of the image. CBIR [7] provides the query like sketch/image which is used to find similar image of same modality, anatomical region and disease associated with the text. Multiple feature representation provides each image information which is expressed by various feature integrated with the neighbor and similar images. An optimal feature is the most difficult problem to computational and data retrieval .another issue in CBIR system is searching and ranking images based on their visual similarity.

The remainder of this paper is organized as follows: Section 2 gives the basic concept involved in the image retrieval system and common client/server system communication model. Section 3 contains the literature review about existing retrieval system and its algorithm to determine the problem description. Section 4 contains the proposed model to provides the solution of the problem and its experimental result

II. FEATURE EXTRACTION

A. CBIR Classification

Normally medical [6] image classified into low level image and high level image which is distinguish based on the content used to store in the clinical database system. High level images possible to retrieve from main block and its subtype (microscopic) based on resolution, contrast and signal to noise ratio. It contains more inner detail of the image for example tissues and its cell types. A low level image contains only basic information which is not concentrate about micro-descriptors. Example: Google images.

B. Feature Selection

Feature selection applies based on the following characteristics of stored medical image content .basically two algorithms used for feature selection greedy algorithm and genetic algorithm. Both cases objective of the function is to maximization of the number of

correctly retrieved images. Greedy algorithm used only less than 10 feature selection. Chromosomes are numbered as an integer in genetic algorithm and it provides the better performance than the greedy algorithm. Some of the image characteristics define the feature selection.

a. Texture Feature

Texture [7] is an important feature of natural image. Different types of techniques developed for measuring texture similarity .some of the text features are degree of contrast, coarseness, directionality, regularity and periodicity. Basically text feature analysis performed by Gabor Filter and Fractal Filter. Gabor filter widely adopted for medical image retrieval.

b. Color Feature

Color [4] is the most efficient visual feature in image retrieval. Color feature retrieval based on primitive color feature that is RGB (Red, Green, and Blue).color is the property which depends on the reflection of light to the human eye. Color feature assigns the each pixel of the image.RGB can be transfer into other color space.

C. CBIR Performance Measures

Content based image retrieval system [8] performance measure by Precision and Recall values. Precision define as the ability of system can retrieved only model that relevant to the given query .Recall define as the ability to retrieved all model that relevant to the given query. Precision and recall used to analysis CBIR system accuracy depends on the size of the clinical database. Ideal CBIR system precision value is equal to one (i.e.) no of relevant image retrieved is equal to total no of image retrieved in the database

$$\text{Precision} = \frac{\text{No of relevant image retrieved}}{\text{No of image retrieved}} \quad (1)$$

$$\text{Recall} = \frac{\text{No of relevant image retrieved}}{\text{No of relevant image in the database}} \quad (2)$$

In real time environment, precision value gradually decreased when the size of the database increased and no of feature selection value is reduced. Recall used to measure the retrieval efficiency of the clinical database. Content based image retrieval system early developed for low feature extraction. But recently microscopic image retrieval introduced for medical image analysis with main disease and its sub type. Every day millions of medical images produced by medical centre so the database system should capable to store those images and retrieve as per pathology requirement.

D. System Overview

System [10] should capable to process tera bytes of data. XML specification for integrating other local and Global information systems (resources).

The primary goal of a graphical user interface (GUI) design ease of use for both first-time and experienced users. All requests can be input with a mouse. No keyboard typing is required. Microscopic medical image content is especially important in that it is the foundation upon which Pathologists make their diagnoses

Table 1. The Client/Server Communication model

Client System	Server System
Apply query to remote server	Assign unique id to related query
Request search service with search parameter	CBIR based search in the database as per given Query
Accept the CBIR based result	Produce full size result Image

E. Retrieval Algorithm

Four primitive image feature types [7] were used to define the content of microscopic pathology images. Each image was processed to obtain color histogram, texture representation, Fourier transformation coefficients, and wavelet coefficients. Recently many algorithms developed for image retrieval especially medical image traditional statistical evaluation methods applied when sample size is small and where domain knowledge is important..An earlier 1990's genetic algorithm used image retrieval. Decision based binary tree algorithm is using content based image retrieval system.

III. RELATED WORK

The authors [2] introduced in this paper Pattern for next generation database system (PANDA) which is involves block based low-level feature extraction. Expectation maximization algorithm used to automatically determine the number of clusters and similarity between the two clusters estimated. Pattern similarity used to retrieve the medical image .Three steps involves in PANDA method, Pattern type and pattern finally class. Pattern type is described about pattern structure. Pattern provides corresponding pattern type and class is a collection of semantically related pattern of the same pattern type. Author tested 6000 image with system configuration 1GB RAM and 60GB hard disk. Experiment result shows the retrieval time reduced 0.22 ±0.04 s.

The authors of [11] and [8] have discussed about textural classification techniques used to retrieve the medical image with improved precision rate .sometime query or feature selection is not meet the user expectation so textural classification which is used to modify the query image as per the content of medical image in the clinical database system. Experiment result shows the half of the computational time used to search entire database and also 10%−15% improvement in precision at each recall level.

Iqbal [10] proposed content based retrieval system that is I-Browse system which combine low-level image processing techniques with high-level semantic analysis of medical image. This method generates the textual annotation especially for the unknown images. I-Browse system also support the natural query language example is Google engine. Browse architecture

functioning with three type of detector 1. Course detector 2.semifine detector 3.fine detector which provides unknown image content retrieval, sub-image retrieval for I-Browse and historical based image retrieval respectively.

Hsu [1] discussed about the high level image retrieval with text based query. Representation of image features like histogram which corresponding to region or sub-image. Self supporting map algorithm introduced for join querying image search scheme. If anything need to change in query keyword before selecting relevance image for the query as well as updated queries are further used to adaptively create a new answer for the next set of returned images as per user preferences .it utilize the search engine to reduce the retrieval time to apply join query method. It efficiently merged the textual and image feature.

Hanka [4] presented about content based image retrieval method which integrates with picture archiving and communication system .it provides the pixel based image retrieval with structured local binary haar pattern and haar wavelet. Haar function defines the color image with basic three colors RGB in each pixel value so based on the pixel value it retrieves the image from clinical database.

Two authors [11], [12] discussed weighted feature in content based image retrieval system .color and texture always important parameter in image retrieval. The performance of image retrieval system enhanced by Combining two feature color and texture with weights so it provides fast image retrieval from the clinical database when compare to all other method because high weights assign important feature which easily identify the large database.

Hatice [14] defines gives solution for multiple image queries for analyzing retrieved image with its subtype. It performed by multitier approach which produced first rank result for microscopic medical image retrieval.

Mohd. Anjum, Md. Asr [15] describe about the rank methodology based on the properties of the image.It perform the priority of weighting value for each pixel as the result of image classification complexity reduced

Anamika Ahirwar [16] explain about the region based classification and also segmentation which is test with the the MRI images.fuzzy neuro techniques used for train the given MRI images

A. Problem Description

Microscopic image classification and discrimination (sub- type) is the most difficult problem in medical image retrieval system. In real time case only few systems has been developed and used in clinical environment.multiimage query is very complex to separate the primary disease type and secondary disease type (i.e.) classification of different type of tissue images (first tier) and matching the given query images (second tier) with classified tissue images (first tier).

Due to increasing large number of medical images in the database, Retrieval image grouping delay which is also increasing. From the literature survey, we find most of the existing system contains average image classification accuracy is the greatest challenge in microscopic images.

IV. PROPOSED SYSTEM

A. Multi-Tier Retrieval Approach for Multi-Image Quries

In this system having the two tier approach where major disease type such as follicular- lymphoma and neuro- plastoma tissue images classified from support vector machine and second tier where query related images retrieved from the database. second tier approach, we will use our proposed multi-image. query and retrieval methodology to retrieve the images from the reference database in the order of their image-level visual matching similarities by preserving the slide-level semantic similarity.

a. First-Tier: Disease Type Classified with SVM

Support vector machine used to classified main disease type such as NB and FL.svm classifier is the traditional method which is used for the content based microscopic image retrieval system .main purpose of svm decision hyper plane for binary classification issue by maximize the margin which is the distance between the hyper plane and the closest data points of each class in the training dataset that are called the support vectors. The hyper plane chooses from all the possible hyper planes through a complex combinational problem optimization so that it maximizes the distance that is the margin.

In our SVM classifier system, we chooses the radial basis function that is one of the most often used kernels and it provides good results than other kernels for the categorization of our data. LibSVM MATLAB code [16] was used.

Recall from the video lectures that SVM classification solves the following optimization problem:

$$\min_{(w,b)} \|\omega\|^2 + C\sum_i^m \varepsilon_i,$$

$$y^{(i)}\left(w^T x^{(i)} + b\right) \geq 1 - \varepsilon_i, i = 1, 2, \cdots, m; \qquad (3)$$

$$\varepsilon_i \geq 0, i = 1, 2, \cdots, m$$

After solving, the SVM classifier predicts "1" If

$$\omega^T x + b \geq 0 \qquad (4)$$

and "-1"otherwise.the
Decision Boundary is given by the line

$$\omega^T x + b = 0 \qquad (5)$$

b. Second-Tier: Image Retrieval with SS-Tree algorithm

An SS-tree is built by inserting vector one by one an initially empty SS-tree .A vector is inserted into a leaf node which is decided by descending from root or selected node .if the leaf node is full ,it is split and a new leaf node is created .the vector in the original leaf

node and vector has been inserted then it can be distributed among the two node.

SS-Tree algorithm outlined below

Insert(r: node, k: entry)
//insert k into a tree rooted at r n=r //start at the root
while n is not leaf node or
at the same level as k   // descending determine a child of n,c ,such that
dist(c,k)=min(dist(c,k)),Ci is a child of n c implies to to n
if n has an empty entry add k into n
change the information in its order else
n'=split(n,p) // n' is the new node
change the information in its order with n'
create entry e to denote n'
insert(r,e)   //re-insert the new node
in this algorithm used to for both leaf and non-leaf entries
.leaf entries contains the input vector but non-leaf entries contains only corresponding node values .related query image can be retrieve from the database by grouping the related features
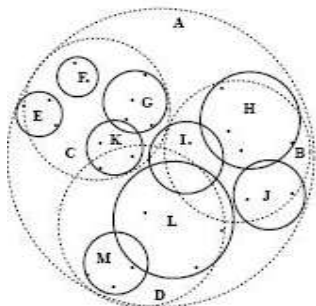
Fig. 1. 2D-Feature Representation

SS-Tree Classified the unique features and grouping the related feature which is reduced the retrieval time and complexity.
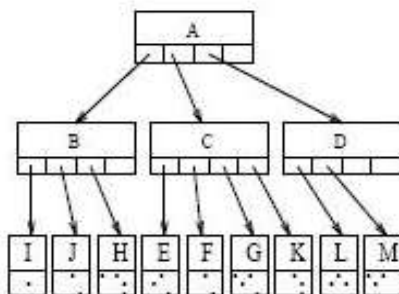

Fig. 2. Representation of SS-Tree

B. Dataset and Experimental Results

Fig. 3 shows randomly selected sample images belonging to different histological grades of FL cases. The number of cropped images per slide is between 11 and 30 for FL cases and between 7 and 35 for NB cases. For FL slides, a team of experienced hematopatho-

logists selected about 10 random microscopic high power field (HPF) to interpret the disease grade in terms of the average number of centroblasts per HPF.
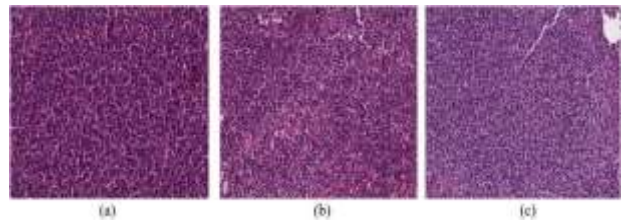

Fig. 3. Sample H&E- stained FL images associated with the three grades (a)Grade I. (b) Grade II. (c) Grade III

$$sensitivity = \frac{TP}{TP + FP}$$

In this equation TP is true positive, FP is false positive. In classification systems, accuracy is often used to evaluate the systems. Accuracy is the overall correctness of the system and is calculated as the sum of correct classifications divided by the total number of classifications.
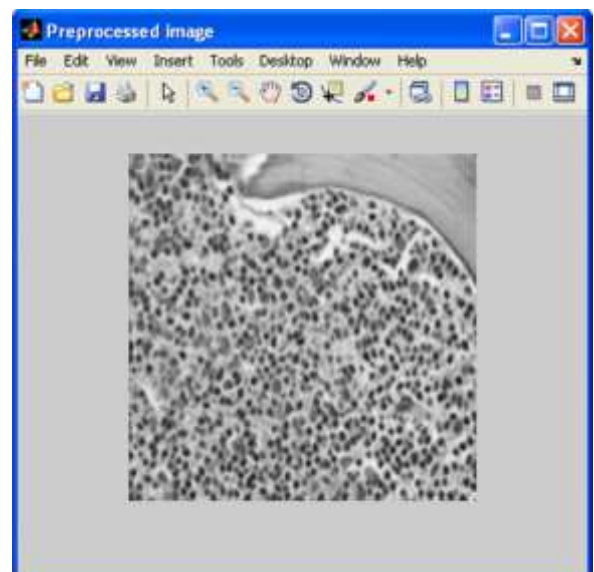

Fig. 4. Input Image

Fig. 5. Proceed Image

Table 2. verage Classification Results (%)

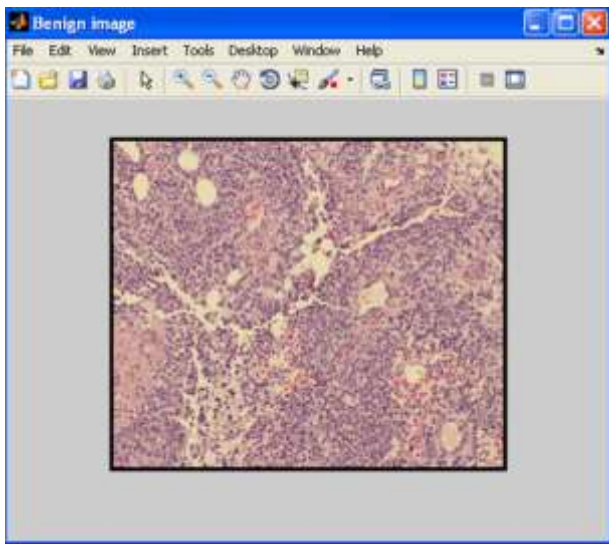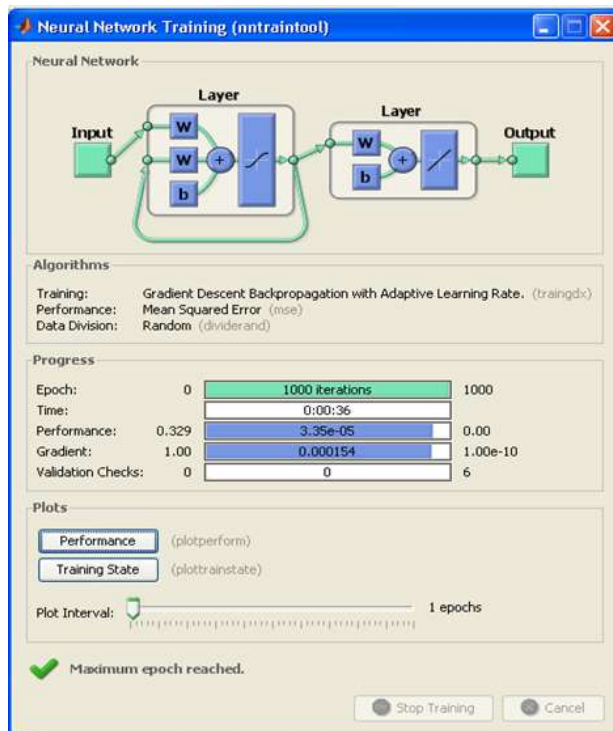| Image Type	No of images	Precision	Sensitivity
FL	40	93	92.5
NB	40	95	94

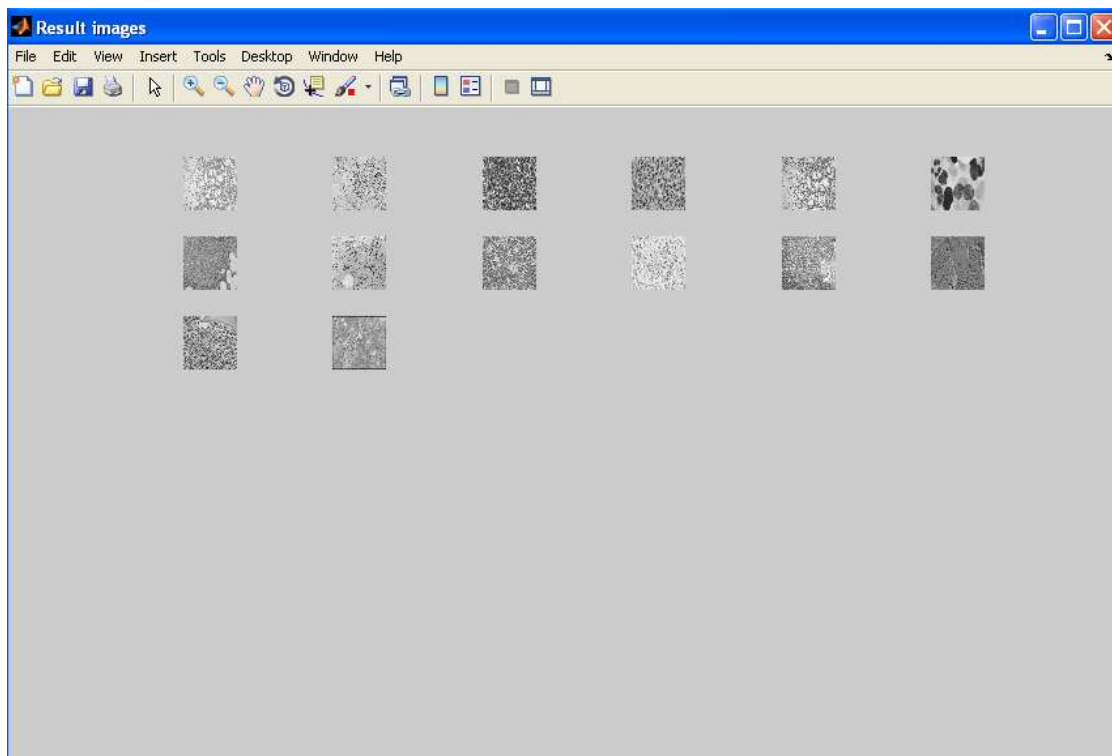Fig. 6. Matching Process

Fig. 7. Training Process

Fig. 8. Output Images

V. CONCLUSION

Results reflect that the proposed system by applying multiple features along with optimization, classification and similarity measurements provides the improved accuracy of the microscopic medical image retrieval system. it's helpful for physician to get more confidence in their decisions for diagnosis and medical research students are get the required images successfully for further investigation of their research. Application of the proposed weighting strategy, inspired by the IR theory which is not limited to microscopic images ,it can be also useful for all type of multiuser searching system and content-based retrieval systems.

REFERENCES

[1] W. Hsu, L. R. Long, and S. Antani, "Spirs: A framework for content- based image retrieval from large biomedical databases," Stud. Health Technol.Informat., vol. 129, no. Pt. 1, pp. 188–192, 2007.

[2] D. Iakovidis, N. Pelekis, E. Kotsifakos, I. Kopanakis, H. Karanikas, and Y. Theodoridis, "A pattern similarity scheme formedical image retrieval", IEEE Trans. Inf. Technol. Biomed., vol. 13, no. 4, pp. 442–450, Jul 2009.

[3] L. Yang, O. Tuzel,W. Chen, P.Meer, G. Salaru, L. Goodell, and D. Foran, "Pathminer:Aweb-based tool for computer-assisted diagnostics in pathology," IEEE Trans. Inf. Technol. Biomed., vol. 13, no. 3, pp. 291-229,MAY 2009.

[4] H. L. Tang, R. Hanka, and H. H. S. Ip, "Histological image retrieval based on semantic content analysis," IEEE Trans. Inf. Technol. Biomed vol. 7,no. 1, pp. 26–36, Mar. 2009.

[5] J. Luo and D. Crandall, "Color object detection using spatial-color joint probability functions," IEEE Trans. Image Process., vol. 15, no. 6,pp. 1443–1453, Jun. 2006.

[6] C.-R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. M. Aisen, andL. S. Broderick, "Assert: A physician-in-the-loop content-based retrieval system for HRCT image databases," Comput. Vis. Image Understand.,vol. 75, no. 1–2, pp. 111–132, 1999.

[7] D. Iakovidis, N. Pelekis, E. Kotsifakos, I. Kopanakis, H. Karanikas, andY. Theodoridis, "A pattern similarity scheme formedical image retrieval," IEEE Trans. Inf. Technol. Biomed., vol. 13, no. 4, pp. 442-450, Jul. 2009.

[8] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," IEEE Trans. Syst., Man Cybern., vol. 3, no. 6,pp 610–621, Nov. 1973.

[9] M. M. Rahman, P. Bhattacharya, and B. C. Desai, "A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback," IEEE Trans. Inf Technol.Biomed., vol. 11, no. 1, pp. 58–69, Jan. 2007.

[10] M. M. Rahman, P. Bhattacharya, and B. C. Desai, "A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback," IEEE Trans. Inf Technol.Biomed., vol. 11, no. 1, pp. 58–69, Jan. 2007.

[11] M.Henning Muller, M.Nicolas Michoux and B.David Bandon Geissbuhler"A Review of Content Based Image Retrieval Application- Clinical Benefits And Future Directions",in Division for Medical Informatics, University Hospital of Geneva 2nd vol published in http://mirc.rsna.org.

[12] M.Mahmudur Rahman, Sameer K.Antani, and George R Thoma,"Alearning based similarity fusion and filtering approach for biomedical image retrieval Using SVM Classification and Relevance Feedback",IEEE Trans on information technology in biomedicine, vol. 15, no. 4, july 2011.

[13] Behjat Siddiquie1 Rogerio S. Feris2 Larry S. Davis1, "Image Ranking and Retrieval based on Multi-Attribute Queries", 2011 IEEE Conference Computer Vision and Pattern Recognition (CVPR), on 20- 25 JUNE 2009.

[14] Hatice Cinar Akakin and Metin N. Gurcan, Senior Member, IEEE Content-Based Microscopic Image Retrieval Systemfor Multi-Image Queries" IEEE Transactions on information technology in biomedicine, vol. 16, no. 4, july 2012.

[15] Mohd. Anjum, Md. Asraful Haque, Nesar Ahmad," Analysis and Ranking of Software Reliability Models Based on Weighted Criteria Value", I.J. Information Technology and Computer Science, 2013, 02, 1-14 Published Online January 2013 in MECS (http://www.mecs-press.org/)

[16] Anamika Ahirwar," Study of Techniques used for Medical Image Segmentation and Computation of Statistical Test for Region Classification of Brain MRI," I.J. Information Technology and Computer Science, 2013, 05, 44-53

Authors' Profiles

P.Suresh completed his B.E, Electronics Engg and M.E Communication Engineering From Muthayammal Engineering College , M.Kumarasamy College of Engineering, Anna university , Madras . Tamilnadu . He is having 2 years Industry experience in database management systems. His area of interested includes image processing, Advanced database systems.

L.Malliga received diploma, Electronics and Communications Engineering from AyyaNadar Janakiammal Women's college,Sivakasi in 1986.After the long break, she completed her AMIE, Electronics and Communications Engineering in 2006, the M.E Applied Electronics from RVS College of Engineering Dindigul in 2009.When she is currently pursuing the Ph.D(Part Time) in the Department of ECE and nearly twenty years of Teaching experience in various Polytechnic colleges and ITI's.Now she is a Senior Assistant Professor in M.Kumarasamy college of Engineering,Karur.

She has published several National and Inter National conferences. She is a Life member of ISTE.Her major fields of interests are Content based Image Retrieval,Bio Medical Imaging and Wavelet Transform.

M.vijay completed his B.E, Computer Science and Engineering and M.E Communication Engineering From Annamalai University, M.Kumarasamy College of Engineering, Anna University, Madras. Tamilnadu . He is having 2 years Industry experience in database management systems. His area of interested includes image processing, Ad-hoc Networks