

A SURVEY ON CONTENT BASED MEDICAL IMAGE RETRIEVAL FOR MRI BRAIN IMAGES

G.Sharmila¹, D.Saraswathi², R.Ramkumar³

¹PG Scholar, ECE, Manakula Vinayagar Institute of Technology, Puducherry, India

²Assistant Professor, ECE, Manakula Vinayagar Institute of Technology, Puducherry, India

³PG Scholar, ECE, Manakula Vinayagar Institute of Technology, Puducherry, India

Abstract

BRAIN tumor is also a common curable disease, when it is diagnosed at the earlier stage. Diagnose of disease is depend 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. In this paper, we present an overview of CBMIR technique to classify brain images. This paper also includes literature survey on feature extraction and feature selection method with similarity matching algorithm. The main objective of this paper is to give comparison on existing methods and highlight the best method to retrieve similar tumor image from a large database.

Keywords: Content Based Medical Image Retrieval (CBMIR), feature extraction, feature selection, similarity matching, magnetic resonance images (MRI), brain tumor images.

1. INTRODUCTION

Content based image retrieval plays an important role in many applications especially in the medical field. Because of thousands of images produced by the hospitals everyday labeling the image and classifying abnormality by human intervention is a time consuming and difficult task. Further treatment planning is depend on how accurately the abnormality is detected, false detection may also cause wrong diagnosis leads to serious problem to face this situation the best content based image retrieval method is used to retrieve query based image from large database.

Accurate result can be produced only with the help of automated computer aided technique. CBMIR can be useful for many diseases such as brain tumor, breast cancer, spine disorder problem etc which is acquired through many modalities such as CT scan, MRI, mammogram etc. In this paper we concentrate on MRI brain tumor 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 classification of brain tumor images is an important part of query based image retrieval from a large database. 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 and also we can't assure that all the content feature does not contain the useful information so the redundant features are reduced using some of feature selection method. Hopefully, this survey will help the medical Practitioner to detect the disease accurately for further treatment and also gives idea for innovation of hybrid technique for research scholar those who work in this area.

This paper gives an overview of available literature in the field of content-based access to medical image data and on the technologies used in the field. Section II provides literature survey on image pre-processing. Section III contains the details of extracted features from an image. Section IV presents the literature survey on feature extraction techniques. Section V presents the literature survey on feature selection techniques. Section VI gives a comparative analysis on similarity matching and classification techniques. Section VII summarizes several concluding remarks and future work.

Most of these systems have a very similar architecture for browsing and archiving/indexing images comprising tools for the extraction of visual features, for the storage and efficient retrieval of these features, for distance measurements or similarity calculation and a type of graphical user interface (GUI). This general system setup is shown in Fig. 1.

2. LITERATURE SURVEY ON IMAGE PRE-PROCESSING

Image preprocessing is the first step of image retrieval to ensure accuracy of subsequent steps. The images acquired through different modalities cause many artifacts such as low resolution, noise and extra cranial tissue etc. which reduces the accuracy of acquired result.

In order to overcome the above problem preprocessing of an image is required. An analysis on filtering techniques such as Gabor & QMF filters for noise is performed by [1]. These primitive methods along with reducing the noise blur the important and detailed structure necessary for subsequent steps. In order to increase the processing speed and to reduce the error probability of mammogram images Morphological top hat filtering algorithm is utilized in [2]. But these types of filtering are applicable for mammogram images only. To eliminate the noise in images, Gaussian filter is suggested in [3].

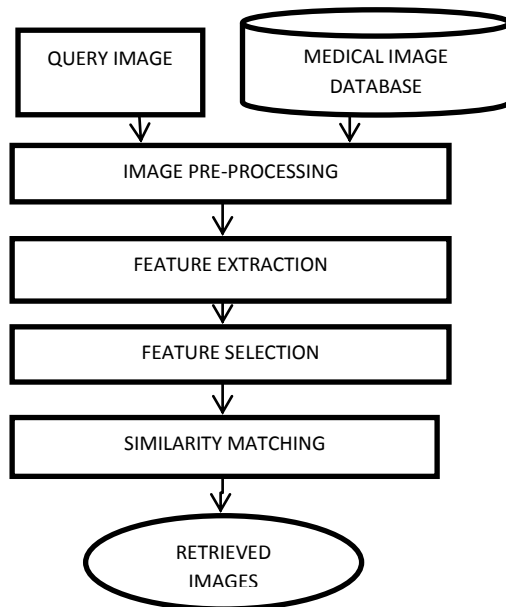


Fig -1: A framework of Content Based Medical Image Retrieval System

The advantage of this filter is to reduce the noise as well to increase the contrast and intensity of an image. To improve image quality histogram equalization, edge detection, noise filtering and thresholding is proposed in [4]. Diffusion filtering combined with simple non-adaptive intensity thresholding is used to enhance the region of interest in [5]. The main drawback of this technique is non-adaptive nature of the threshold value. The noise removal technique using wavelet and curve let is implemented in [6]. Hybrid approach involving variance stabilizing transform (VST) are also used. But this technique is applicable only for image with Poisson noise. Combination of approaches such as local maximum,

local minimum, local $(\max - \min)/2$, local mean, local median, local Mod are given in [7]. This will improve the image information content by suppressing the undesired distortions. A contrast agent accumulation model is implemented in [8]. This improves only the contrast characteristics of an image but it is not eliminating the unwanted tissue. Weiner filter is used to remove noise in MR brain image is used in [9]. Apart from noise removal several other preprocessing steps are also to be considered in CBMIR. This includes image format conversion, image type conversion etc. the combination of three modalities of MR images for further processing is proposed in [10]. Though several preprocessing method are mentioned for removal of noise in an image, the removal of unwanted tissue and preserving needed image information content without blurring is highly essential for accurate retrieval of image from a large database. Some of the Pre-processing techniques are tabulated below:

Table -1 Comparison of various preprocessing methods

METHODS	CHARACTERISTICS	DISADVANTAGES
Wavelet Transform [6]	Analysis image in different frequencies	Missing of some original information during decomposition.
Thresholding [4]	Enhance the region of interest	Non-Adaptive nature of threshold value
Gabor and QMF [1]	Removes only the noise	Blur the details present in an image
Morphological operator	Used to segment the region of interest	Doesn't remove noise or improve image quality

3. TYPES OF FEATURES

3.1 Visual Features

This section describes the visual features that are used in the various applications. Feature contains information about image content that is needed to retrieve similar query images from the database and also to classify the type of disease in the medical field. Visual features were classified in [12] into *primitive* features such as color or shape, *logical* features such as identity of objects shown and *abstract* features such as significance of scenes depicted. Still, all currently available systems only use primitive features unless manual annotation is coupled with the visual features as in [13].

3.2 Color Features

In stock photography (large, varied databases for being used by artists, advertisers and journalists), color has been the most effective feature and almost all systems employ colors. Although most of the images are in the red, green, blue (RGB) color space, this space is only rarely used for indexing and querying as it does not correspond well to the human color perception. Much effort has also been spent on creating color spaces that are optimal with respect to lighting conditions or that are invariant to shades and other influences such as viewing position [14, 15]. This allows identifying the colors even under varying conditions but on the other hand information about the absolute colors is lost. But in the medical domain, absolute color or grey level features are often of very limited expressive power unless exact reference points exist as it is the case for computed tomography images.

3.3 Shape Features

Shape describes the edge of an image however result of segmentation is also a one of the shape feature. Pattern recognition based on color and shape is mentioned in [22]. Fast and effective retrieval of tumor shape is given in [23]. For patients with poor MRI quality, the texture and intensity features may prove inadequate for PF tumor segmentation. Though it gives information about images, the internal tissue present in a tumor image may be lost and it causes misclassification.

3.4 Texture Features

Due to complex tissues present in an MR brain images classification through shape feature is difficult. For accurate classification, features that are used in medical field is called texture. Texture is a commonly used feature in the analysis and interpretation of images. It is characterized by a set of local statistical properties of pixel intensities. Analysis of texture feature is mentioned in [7] which improve the accuracy of classification.

3.5 Features used in Medical Field

As color and grey level features are of less importance in medical images than in stock photography, the texture and shape features gain more importance in [26]. Basically all of the standard techniques for texture characterization are used from edge detection using Canny operators [16] to Sobel descriptors [17]. [18,21,19] also use Fourier descriptors to characterize shapes, [16,20,19] use invariant moments and [16] also scale-space filtering.

4. LITERATURE SURVEY ON FEATURE EXTRACTION

The intention of Feature extraction is to reduce the original data set by measuring certain properties, or features, that

distinguish one input pattern from another pattern [25]. The extraction of feature vector which consists of various feature components. It is generated to represent the content of each image in the database with accuracy and uniqueness. Spatial gray level co-occurrence matrix estimate the image properties related to second order statistics. Haarlick [32] suggested the use of GLCM is one of the most well known method. The wavelet based texture feature for classification is used by [28]. Multi-fractional Brownian motion (MBM) algorithm is used in [27]. The advantage of this method is image with different resolution gives same result. Modified Haar wavelet transformation is proposed in [29]. The texture features namely Contrast, correlation Homogeneity and Energy is used in [30]. This improves specificity and accuracy of retrieved image. Auto color correlogram and correlation in [31] will get accuracy in less iteration. But the iteration is depending on the need of application. EI-Sayed et al. [24] has obtained the features related with MR images using discrete wavelet transformation (DWT). But all the techniques which use basic discrete wavelet transform which does not yield superior result. An improved version based on wavelet packet decomposition is implemented in [32]. The results revealed that the extracted features are efficient than previous method. From the above mentioned techniques, it is clear that the extraction of appropriate features will improve the accuracy for classification and similarity matching.

Table- 2 Mean square Error and classification accuracy of different feature extraction technique

Stages	Mean Square Error			Classification Accuracy%		
	GL CM	HAR ALIC K'S	WSF TA	GLC M	HAR ALIC K'S	WSF TA
Trainin g	1.596	2.235	0.0585	98.6%	89%	100%
validati on	2.345	8.632	1.203	98.7%	87.6%	100%
Testing	1.895	2.2154	1.003	96.4%	88.9%	98.0%

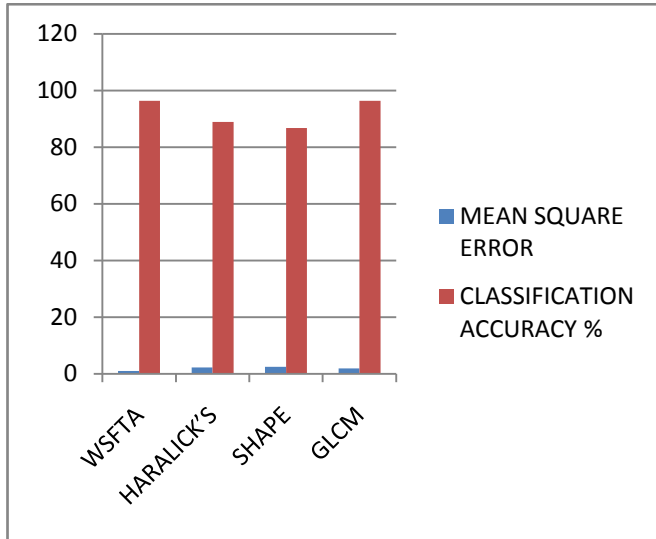


Chart -1: performance of different feature extraction technique

5. LITERATURE SURVEY ON FEATURE SELECTION

Feature selection is the process of reducing the dimension of feature vectors. For feature selection, every feature is observed. Significant features are selected by the calculation of mean values for every feature in benign tumour class and malignant tumour class. Student's t-test [34] is employed as a method to separate both the classes. Principal component based analysis is used in [33] gives better feature subset, but fresh training is needed whenever it encounter a new images. The ability of genetic algorithm for pattern classification is explored in [34]. The samples belonging to the same class is only accepted by GE. Reduction of features using Linear Discriminant Analysis is proposed in [35].

An another evolutionary approach namely, Partial swam intelligence is used by [36]. This method is used for optimizing the weight of Artificial Neural Network for image classification application. Since PSO involves many random parameters, the weight obtained through this may not be a stabilized set of weight. Ant Colony based feature selection is given in [37]. By selecting the features in terms of feature set size, its computation increases when feature size increases. Optimal Fuzzy rule selection for classification is implemented in [38]. But this technique is highly sensitive to change in the parameters of the membership functions. The hybrid of above techniques will give even better results for eg: (wavelet+ Principal component analysis) or (genetic algorithm+ fuzzy rule) etc is mentioned in [39]. But time complexity is more in the hybridized method which makes system non- feasible. The above mentioned feature selection methods can be used for calculating classification accuracy and similarity matching.

6. LITERATURE SURVEY ON CLASSIFICATION

Automatic similarity matching is used to retrieve similar images from large database that reduces the cost of classification work of human. One of the primary distance measure techniques is Euclidean distance which is used in [39]. The dimensionality is often reduced to work efficiently with these distances even in large database. The neural network is used for classification of mammogram images [40]. Other statistical approach Bayesian network is used in [42]. Classification using Support Vector Machine is mentioned in [41] for MR brain images which is the simplest one and gives better accuracy. The Hybrid technique of SVM and probabilistic neural network (PNN) classifiers is proposed in [34]. The above mentioned techniques gives an improved accuracy for classification and image retrieval.

7. CONCLUSIONS AND FUTURE WORK

We have categorically presented a detailed survey of the prominent methods of automated technique for medical image retrieval published in the literature. In the case of Pre-processing stage, various types of filter techniques have been discussed. In the case of feature extraction and feature selection some of the evolutionary techniques used in the literature have also been presented. The performance depends on the appropriate extracted features and the features that are selected for similarity matching [46]. We hence presented a survey on all stages as a whole to implement the improvised CBMIR system, which has a rich potential and scope of application in medical domain. CBMIR is an ever- growing rich field, with a wide scope of interdisciplinary research over the fields of Biomedical, Computer Science, Mathematics, Statistics and Machine learning.

Table-3: Classification accuracy of different metrics

TECHNIQUES	CLASSIFICATION ACCURACY
Minkowski	79.40
Chebyshev	79.20
L2	79.80
Standardized L2	79.20
Mahalanobis distance	80.80
City block	81.40
Correlation	81.60
Relative deviation	81.00
L1	81.60
cosine	82.00
Spearman	82.40

In the future, one can expect rich concepts from widely varying areas such as information theory, optimization theory, rough-fuzzy sets, soft computing and evolutionary techniques etc., to contribute and enrich this problem domain in the field of pattern recognition.

REFERENCES

- [1] Sebe N, Michel S.L “Wavelet based texture classification”. 15th Int. conf. on pattern recognition 2000, Vol 3: 947-50.
- [2] D. Cascio, F. Fauci, R. Magro, G. Raso, “Mammogram Segmentation by Contour Searching and Mass Lesions Classification With Neural Network” IEEE Transactions On Nuclear Science, Vol. 53, No. 5, October 2006.
- [3] Nitish Zulpe and Vrushsen Pawar, “GLCM Textural Features for Brain Tumor Classification” International Journal of Computer Science Issues, Vol. 9, Issue 3, No 3, May 2012.
- [4] Ms.suchithra swami, “Brain tumor detection using unsupervised learning based neural network” Int. conf on comm system and network technologies, IEEE 2013.
- [5] Yang Y, Hung S, “Novel statistical approach for segmenting of brain magnetic resonance imaging using an improved expectation maximization algorithm” Optica Applicata 2006 ;36:125-36.
- [6] Zhang B, Jalal M, Strack J. “ Wavelet , Ridgelets and curvelets for poisson noise removal” IEEE Trans.Image .proc.2008;17:1093-1108.
- [7] B.V. Ramana Reddy et.al “ Classification of Textures Based on Features Extracted from Preprocessing Images on Random Windows” Int. journal of Advanced Science and technology 2009; vol 9.
- [8] Prastawa M, bullit E, Gerig G. “simulation of brain tumors in MR images for evaluation of segmentation efficiency” Medical image analysis 2009; 13: 297-311.
- [9] Ratan R, Sharma s, Sharma s. “ Multiparameter segmentation and quantization of brain tumor from MR images”. Int. jrnal of science and technology 2009; 2: 11-5.
- [10] Ratan R, Sharma S ,Sharma S. “Brain tumor detection based on multi-parameter MR image analysis”. Graphics, vision and image processing Journal 2009; 9: 9-17.
- [11] D. Judehemanth, C. keziselvavijila, J. anitha ‘Artificial intelligent based Brain Pathology Identification techniques in Magnetic resonance Images”. Int.jrnal of review in computing. 2009.
- [12] J.P. Eakins, M.E. Graham, “Content-based image retrieval” ,Tech. Rep. JTAP-039, JISC Technology Application Program Newcastle upon Tyne, 2000.
- [13] T. Pfund, S. Marchand-Maillet, “Dynamic multimedia annotation tool” Proceedings of the SPIE Photonics West Conference on Internet Imaging III, vol. 4672, San Jose, CA, USA, 2002, pp. 216—224
- [14] J.-M. Geusebroek, R. van den Boogaard, A.W.M. Smeulders, H. Geerts, “Color invariance” , IEEE Trans. PatternAnal. Machine Intel.23 (12) (2001) 1338—1350.
- [15] M. Ortega, Y. Rui, K. Chakrabarti, K. Porkaew, S. Mehrotra,T.S. Huang, “Supporting ranked boolean similarity queries in MARS”, IEEE Trans. Knowledge Data Eng. 10 (6) (1998) 905—925.
- [16] K. Veropoulos, C. Campbell, G. Learnmonth, “Image processing and neural computing used in the diagnosis of tuberculosis”, Proceedings of the Colloquium on Intelligent Methods in Healthcare and Medical Applications (IMHMA), York, UK, 1998.
- [17] C. Brodley, A. Kak, C. Shyu, J. Dy, L. Broderick, A.M. Aisen, “Content-based retrieval from medical image databases: A synergy of human interaction, machine learning and computer vision” ,Proceedings of the 10th National Conference on Artificial Intelligence, Orlando, FL, USA, 1999, pp. 760—767.
- [18] S. Antani, L.R. Long, G.R. Thoma, “A biomedical information system for combined content-based retrieval of spine X-ray images and associated text information”, Proceedings of the Third Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP 2002), Ahamdabad, India, 2002.
- [19] J.M. Bueno, F. Chino, A.J.M. Traina, C.J. Traina, P.M. Azevedo-Marques, “How to add content-based image retrieval capacity into a PACS”, Proceedings of the IEEE Symposium on Computer-Based Medical Systems (CBMS 2002), Maribor, Slovenia, 2002, pp. 321—326.
- [20] M.E. Mattie, L. Staib, E. Stratmann, H.D. Tagare, J. Duncan, P.L. Miller, PathMaster “ Content-based cell image retrieval using automated feature extraction”, J. Am. Med Informatics Assoc. 7 (2000) 404—415.
- [21] A.K. Jain, A. Vailaya, “Image retrieval using color and shape”, Pattern Recog. 29 (8) (1996) 1233—1244.
- [22] P. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, Z. Protopapas, “Fast and effective retrieval of medical tumor shapes”, IEEE Trans. Knowledge Data Eng. 10 (6) (1998) 889—904.
- [23] E.-S. A EI-Dahshan, T. Hosny, and A-B. M. Salem, "Hybrid intelligent techniques for MRI brain images classification," Digital Signal Processing, vol. 20, no. 2, pp. 433-441, 2010.
- [24] Lawrence O Hall, “ A comparison of neural network and Fuzzy clustering techniques in segmenting magnetic resonance images of the brain”, IEEE Trans on neural networks, Vol 3, No.5,sep 1992.
- [25] Anil. k. jain, “Fundamentals of digital image processing” prentice-hall publication, ISBN: 0-13-336165-9, 1989.
- [26] Henning Müller, “A review of content-based image retrieval systems in medical applications—clinical

- benefits and future directions' International Journal of Medical Informatics (2004) 73, 1-23
- [27] Atiqislam, syed M.S reza and k.m. iftekharuddin "Multi fractal texture estimation for detection and segmentation of brain tumor" *IEEE Trans.on.medical,2013*.
- [28] Arivazhagan S, Ganesan L., " Texture classification using wavelet transform". Pattern recog letters 2003;24: 1513-21.
- [29] P.S. Malge et.al "Performance Evaluation of Texture based Image Retrieval" International Journal of Computer Applications (0975 – 8887) Volume 72– No.2, May 2013
- [30] RamadassSudhir et. Al, " A Efficient Content based Image Retrieval System using GMM and Relevance Feedback", International Journal of Computer Applications (0975 – 8887) Volume72– No.22, June 2013.
- [31] R. M Haarlick, "Statistical and structural approaches to texture", proceeding of IEEE,vol.67, pp.786-804, 1979.
- [32] Hiremath Pet.al "Texture classification using wavelet packet decomposition" Graphics, vision, and image processing journal 2006;6:77-80.
- [33] Kishore K,Patnaik M, Mani V, Agarwal K. "Applications of genetic programming for multi category parttern classification" , IEEE trans.on Evolutionary computation 2009; 4: 242-58.
- [34] A. Padma Nanthagopal, R. Sukanesh, "Wavelet statistical texture features-based segmentation and classification of brain computed tomography images" IET Image Processing 2012.
- [35] V.P.Gladis Pushpa Rathi and Dr.S.Palani "Brain Tumor MRI Image Classification With Feature Selection And Extraction Using Linear Discriminant Analysis" 2012.
- [36] Chandramouli k, Izquierdo E , "Image classification using chaotic particle swam optimization" IEEE Int. conf, on image processing 2006; 3001-04.
- [37] Ling Chen, Bolun Chen, Yixin Chen, " Image Feature Selection Based on Ant Colony Optimization" 2010.
- [38] Chakraborty D,Nikhil P.A, "Neuro fuzzy scheme for simultaneous feature selection and Fuzzy rule based classification" IEEE Trans, neural network 2004; 15: 110-23.
- [39] Aditi Mehta , "Review and Comparison of Various Feature Extraction Techniques in CBIR", International Journal of Computer Applications (0975 – 8887) Volume 71– No.23, June 2013.
- [40] R. Hanka, T.P. Harte, "Curse of dimentionaliry: classifying large multidimensional images with neural network", Proceedings of the European workshop Computer – intensive methods in control and signal processing (CIMCSP 1996), prague, Czech Republic, 1996.
- [41] Magadiet.al, "An intelligent model for automatic brain tumor diagnosis bared on MR images".
- [42] Aditi P. Killedar, " Content Based Image Retrieval Approach to Tumor Detection in Human Brain Using Magnetic Resonance Image" 1st International Conference on Recent Trends in Engineering & Technology, Mar-2012.
- [43] Y. Liu, F. Dellaert, "Classification – driven medical image retrieval", Proceedings of the ARPA image understaning workshop, 1997.
- [44] D.Saraswathi, E.Srinivasan, G.Sharmila, "An Automated Diagnosis system using Wavelet based SFTA Texture Features" IEEE Int conf., ICICES, 2014.
- [45] D.Saraswathi, Chanemouga priya, G.Sharmila, "An Efficient Color Feature Extraction Method for Content Based Image Retrieval", NCWCAPS"13, pg: no: 7- 11,2013.
- [46] D.Saraswathi, E.Srinivasan, G.Sharmila, "A Comprehensive study of Classification Techniques used in Medical Field" ICIEES, Int conf., PSG college of Technology, PP. 233-238, 2013.

BIOGRAPHIES



Sharmila Govindaraj received her B.E. degree in Electronics and communication Engineering from Mailam Engineering College, Anna University, Chennai in 2011 and currently doing her Master degree in Electronics and communication engineering at Manakula Vinayagar Institute of Technology, Pondicherry University, Puducherry. Her fields of interests are Image Processing, Digital Signal Processing and Electronic Devices



Saraswathi Velmurugan received her M.E degree in Applied Electronics from PSG College of Technology, Anna university, Coimbatore in 2009 and currently working as an Assistant Professor at Manakula Vinayagar Institute of technology, Pondicherry University, Puducherry. She is pursuing her Ph.D. from Pondicherry University, Puducherry. She has ten years long experience in the field of teaching. Her research interest includes Image Processing, Medical Image Classification, Image retrieval and Recognition Systems.



Ramkumar Ramachandran received his B.E. degree in Electronics and communication Engineering from Mailam Engineering College, Anna University, Chennai in 2011 and currently doing his Master degree in Electronics and communication engineering at Manakula Vinayagar Institute of Technology, Pondicherry University, Puducherry. His field of interests are Image Processing, Wireless and Telecommunication and Computer Networks.