SVM Based Liver Tumor Classification from Computerized Tomography Images

Priya V., Biju V. G.

Abstract— Accurate liver segmentation and tumor detection on computerized tomography (CT) images is a crucial task in the cases where surrounding tissues have intensities similar to that of the liver and lesions reside at the liver edges. In this paper, an automated method to segment liver portion, followed by tumor area from abdominal CT image is proposed. For this, the CT images are pre-processed by median filter to remove noise from the image and liver is segmented using localized region based active contouring algorithm. Tumor is detected from segmented liver using seed region growing algorithm. Using Grey Level Co-occurrence Matrix (GLCM), the texture features of the tumor are extracted. Support Vector machine (SVM) is used to classify the tumor as either benign or malignant based on these texture features. The performances of liver segmentation and tumor detection are evaluated by using Segmentation Matching Factor (SMF), Dice coefficient (DICE COEFF), Root Mean Square Error (RMSE) and Peak Signal to Noise Ratio (PSNR). Experimental results show that the proposed method has a lower error in segmenting the liver and is able to detect and classify all tumors from the liver accurately

Index Terms—CT image, Localized region based active contours, Region Growing Algorithm, GLCM, SVM.

I. INTRODUCTION

The abnormal and uncontrolled growth of cells causes the development of tumor in an organ. The tumor can be benign (non-cancerous) or malignant (cancerous). Liver with malignant tumor is called liver cancer. Hence it is necessary to detect the types of tumor. In order to determine the tumor, the liver is to be segmented. Since liver segmentation is time consuming when it is done manually, an automatic method has to be used. There are different methods for liver segmentation based on some statistical, atlas, geometric, intensity model. Solar [1] proposes a fully automatic method to segment the liver from contrast enhanced CT images. However this method requires a special protocol in order to implement a phase and hence this is not used for general practice. Lamecker presents an automatic approach for segmentation of the liver from CT images based on a statistical shape model [2]. However in this method, there is no clinical evaluation and the selection of the landmarks is very difficult due to the invariant shape of the liver. The intensity based method requires manually defined threshold values [3]. But in atlas based approaches, multiple local errors occur which cause the inaccurate determination of liver boundary [4].

In addition there were several methods such as region growing active contour level set classification algorithm. In medical image segmentation, the level set method family has been successfully used. The advantages of the level set approach are that it handles topological changes and defines the problem in higher dimensions [5]. The active contour method is used to segment abdominal organs in the clinical practice [6]. It works well on native images, because the organs are homogeneous. It is used to detect objects in image using curve evolution techniques. The region-growing based approaches can provide good results on contrast enhanced images [7] and it can be efficiently close round the vessels and tumors.

There were several methods for tumor detection. In multi stage automatic hepatic tumor detection method proposed by Seo [8], tumor is segmented by using optimal threshold value with minimum probability error. But this method produces inaccurate detection for small tumors. Dipak et al. proposes an automatic method that uses spatial fuzzy C means segmentation technique and flood fill algorithm to detect hepatic lesions from CT scans [9]. Nader et al. proposed a method for detecting the tumor from CT image by using a combination of contrast enhancement and isodata thresholding [10]. Sajith et al. proposed a method for segmenting the liver tumor from CT images. It contain Fuzzy C means algorithm followed by level sets method for segmenting the liver tumor from CT images [11]. In Section II proposed method is discussed. Dataset and parameters used are described in Section III. Section IV describes the results and discussion. Section V describes the conclusion.

II. PROPOSED METHOD

Fig. 1 shows the block diagram of the proposed method.

A. Preprocessing

Median filter is used to remove noise and preserve the edges. In order to increase the accuracy of liver segmentation, abdominal CT images are to be preprocessed.
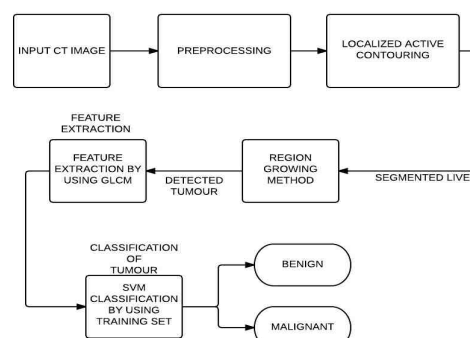

Fig 1: Block diagram of the proposed method

B. Localized Region Based Active Contouring Algorithm

A local region based framework [12] with energy minimization approach is used for guiding active contours and it is used for the liver segmentation. The basic idea of the localized region frame work is that the foreground and background region are defined in terms of local regions. Construction of local energies from each point along the curve is the result of analyzing local regions. For computing local energies the local neighborhoods are divided into local interior and local exterior by the curve.

Let I be the image we want to segment defined on the domain Ω, and let C be a closed contour denoted by the zero level set of a signed distance function $\Phi(x)$. The interior of the contour is represented by the smoothed Heaviside function $H\Phi(x)$ as

$$H\Phi(x) = \begin{cases} 1 & \Phi(x) < -\varepsilon \\ 0 & \Phi(x) > \varepsilon \\ \frac{1}{2}\{1 + \frac{\Phi}{\varepsilon} + \frac{1}{\pi}\sin\frac{\pi\Phi(x)}{\varepsilon}\} & otherwise \end{cases} \quad (1)$$

Where $[-\varepsilon, \varepsilon]$ denote the boundary of the Heaviside function. The exterior of closed contour C is given by the equation $\{1 - H\Phi(x)\}$.

In order to calculate only the local energies, localization radius R is defined in terms of the size of the image and is calculated using

$$R \propto (P_r + P_c) \quad (2)$$

Where P_r is the number of rows in the image and P_c is the number of columns in the image.

Chan–Vese energy is used to model the interior and the exterior of the contour for segmentation. The localized version of the Chan–Vese energy function is given by

$$E_{cv} = (I(y) - \mu_{int})^2 + (I(y) - \mu_{ext})^2$$

(3)

where μ_{int} and μ_{ext} are the interior and exterior intensities of the contour.

C. Seed Region Growing Algorithm

After segmenting liver from the CT image, region growing algorithm is applied to detect the tumor area [13]. It is used to partition the image into a set of connected region, which find the homogeneous regions for fast and good result for tumor detection. It considers two points they are; 1) seed points 2) similarity constraints. Region is grown from the seed point by adding neighbour pixels.

Starting from seed points and region will grow by adding neighborhood pixels that have the properties similar to the seed points. That is, the similarity measure between seed point and each neighborhood pixel is calculated. The pixel with the minimum distance is added to grown region. The process continues iteratively by comparing all neighboring adjacent

pixels of the grown region using the similarity measure. These processes will stops when the intensity difference between region mean and new pixel becomes larger than a certain threshold.

After detecting tumor from liver region, using tumor burden the percentage of total tumor can be calculated.

D. Tumor Burden Calculation

The tumor burden (TB) is defined as the percentage of the total tumor occurs in the liver. Tumor burden is used to analyze the growth of disease, and also helps to find out whether the tumor is cancerous or non-cancerous. It is calculated by using the equation

$$TB = \left(\frac{A_t}{A_l}\right) \times 100 \quad (4)$$

Where A_t is the total tumor area, A_l is the liver area.

Severity of the cancer (S) of liver can be calculated with the help of tumor burden by using the following equation:

$$S = TB \times 10 \quad (5)$$

E. Tumor Feature Extraction Using GLCM

Tumor feature include size, shape, texture and enhancement. Texture is a characteristic of an object. Gray Level Co-occurrence Matrix (GLCM) is used to analysis of tumors [14]. A co-occurrence matrix is used to represent the distance and angular spatial relationship over an image sub region of specific size. GLCM is created from a gray-scale image and is also known as the Gray Level Spatial Dependence Matrix. How many times a pixel with Gray level value i occurs either horizontally, vertically or diagonally to adjacent pixels with value j is calculated by using GLCM.

GLCM direction of analysis is

1) Horizontal direction (0^0)

2) Vertical (90^0)

3) Two type diagonal direction;

a) Bottom left to top right (45^0)

b) Top left to bottom right (135^0)

Some of the texture parameters determined as follows

1) Contrast:

It is the measurement of intensity contrast between a pixel and its neighbouring pixels over the whole images, calculated using the equation:

$$\sum_i \sum_j m_{ij} * (i-j)^2 \quad (6)$$

and it range from 0 to $(size(GLCM,1)-1)^2$

2) Energy

Energy is also known as uniformity. It is defined as the sum of squared elements in the GLCM. It can be calculated by using the equation:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} m_{ij}^2 \quad (7)$$

3) Entropy

It is the measure of randomness, can be calculated by using

the equation:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \log m_{ij} \qquad (8)$$

4) Homogeneity: -

It evaluates the proximity of the distribution of the elements in the GLCM to the GLCM diagonal. It can be calculated by the equations:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{m_{i,j}}{1+(i-j)^2} \qquad (9)$$

5) Mean and Variance

6) Skewness:

It is the measure of symmetry.

7) Kurtosis:

It is the measure of whether the data are peaked or flatted.

8) Cluster shade and Cluster prominence:

They are the measures of skewness of the matrix.
Cluster shade can be calculated using the equation:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} m_{ij} + (i+j-\mu_x-\mu_y)^3 \qquad (10)$$

Cluster prominence can be calculated using the equation:

$$\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} m_{ij} + (i+j-\mu_x-\mu_y)^4 \qquad (11)$$

where $\quad \mu_x = \sum_{i=0}^{N-1} i * m_x(i) \quad$ and $\quad \mu_y = \sum_{j=0}^{N-1} i * m_y(j) \quad$.

F. Classification of Tumor by Using SVM

Support Vector Machines are supervised learning methods used for classification. Here SVM classifies given set of data in to two classes, making it a non-probabilistic binary classifier. From the trained datas, an SVM training algorithm creates a method that assigns new testing datas in to one of two categories. Here the tumors are classified as benign or malignant. SVM builds a hyperplane which can be used for classification. The distance between the separating hyperplane and the nearest points is called the margin of SVM classifier. A good hyperplane is the one that makes the largest margin between two classes. SVMs find the data points near the hyperplane which are called support vectors.

III. DATASET AND PARAMETERS USED

In this paper, abdomen CT images with liver tumor are used as the database. The performance of the segmentation method is calculated by using the parameters such as SMF, DICE COEFFICIENT, PSNR, RMSE etc.

1. Segmentation Matching Factor (SMF)

The accuracy of the segmentation is calculated by using the equation

$$SMF = \frac{A \cap A_{ref}}{A \cup A_{ref}} \times 100 \qquad (12)$$

where A is the segmented output obtained by the algorithm and A_{ref} is the reference image. A perfect match is indicated by a 100% score, a reasonable segmentation is indicated by any score higher than 50%, and any score less than 50% indicates poor segmentation.

2. Dice Coefficient (D)

It is an overlap measure between two binary images. It is used to evaluate the performance of the segmentation and the percentage of D ranges between 0 (no overlap) and 100 (perfect overlap). It is calculated by using the equation

$$D = \frac{2|A \cap A_{ref}|}{|A| + |A_{ref}|} \times 100 \qquad (13)$$

where A is the segmented output obtained by the algorithm and A_{ref} is the reference image.

3. Root Mean Square Error (RMSE)

It can be calculated by the equation

$$RMSE = \sqrt{MSE} \qquad (14)$$

where MSE is calculated by using the equation

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I(i,j) - K(i,j))^2 \qquad (15)$$

where K is the segmented output by the algorithm, I be the reference image.

4. Peak Signal to Noise Ratio (PSNR)

It can be defined as

$$PSNR = 10 \log_{10}\left(\frac{MAX_i}{MSE}\right) \qquad (16)$$

where MAX_i is the maximum possible pixel value of image.

IV. RESULT AND DISCUSSION

The proposed method is tested over two groups of abdomen CT images. The first group contains 20 images. Among these 20 images, 10 images contain benign tumor and other 10 images containing malignant tumor. The second group contains 10 images where the tumor types are unknown. The proposed segmentation method is first applied on first group of images. Fig 2 shows the image of benign type tumor after applying the segmentation method. Fig 2(a) shows the input image. By applying median filter, the input image is

preprocessed and it is shown in Fig 2(b). Localized contouring algorithm is performed to get the segmented liver which is shown in Fig 2(c). After performing seed region growing algorithm to the segmented liver, the detected tumor is shown in Fig 2(d).

Similarly the proposed segmentation method is applied on one of the images of malignant tumor and is shown in Fig 3. Fig 3(a) shows the input image. By applying median filter, the input image is preprocessed and it is shown in Fig 3(b). Localized contouring algorithm is performed to get the segmented liver which is shown in Fig 3(c). After performing seed region growing algorithm to the segmented liver, the detected tumor is shown in Fig 3(d).
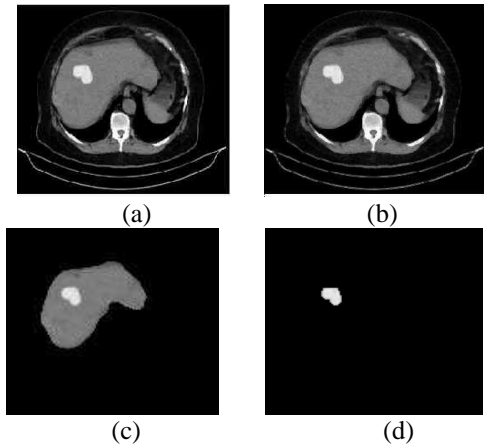

(a) (b)

(c) (d)

Fig.2. (a) Input CT image (b) Preprocessed image(c) Segmented liver (d) Detected tumor.
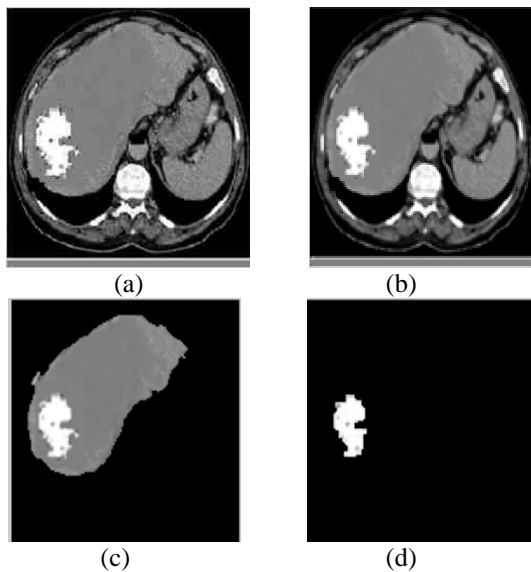

(a) (b)

(c) (d)

Fig.3. (a) Input CT image (b) Preprocessed image (c) Segmented liver (d) Detected tumor.

20 tumors obtained from the first group are taken as the training set of SVM for tumor classification. These 20 tumors are also used as the testing data to see the correctness of classification. It is found that they are correctly classified as either benign or malignant. Then the proposed method is applied on the second group of images in which tumor type is unknown. The tumors thus detected are classified using SVM. The proposed method is applied on one such image and it is shown in Fig 4. Fig 4(a) shows the input image. By applying median filter, the input image is preprocessed and it is shown

in Fig 4(b). Localized contouring algorithm is performed to get the segmented liver which is shown in Fig 4(c). After performing seed region growing algorithm to the segmented liver, the detected tumor is shown in Fig 4(d).

The accuracy of the liver segmentation and tumor detection are done with help of the reference images. The reference images of two group of images are created using ECOGNITION SOFTWARE. After performing an algorithm named 'contrast split segmentation' from this software, the reference image of liver is obtained. Similarly, reference image of tumor is obtained by using 'multi resolution segmentation' from this software.

Accuracy of the liver segmentation method is evaluated by using different parameters with the help of reference image created by ECOGNITION SOFTWARE. The TABLE 1 below shows the segmentation accuracy of some of the images. According to the results shown in the TABLE 1, when comparing the SMF of different source images, the proposed method gives an average of 97.07% accuracy on segmenting liver. The maximum value of DICE COEFF is 99.85%. By comparing RMSE and PSNR the proposed algorithm has a lower RMSE rate of 0.0244 and higher PSNR value of 80.38.
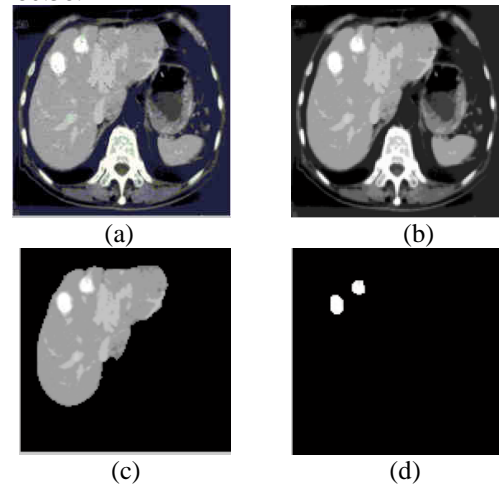

(a) (b)

(c) (d)

Fig.4. (a) Input CT image (b) Preprocessed image (c) Segmented liver (d) Detected tumor.

Accuracy of the tumor detection evaluation is done by using the same parameters as mentioned above and it is shown in the TABLE II. When comparing the SMF of different source images, the proposed method gives an average of 96.35% accuracy. The maximum value of DICE COEFF is 99.85%. By comparing RMSE and PSNR, the proposed method has a lower RMSE rate of 0.0103 and higher PSNR value of 87.84.

Table I. Liver Segmentation Accuracy

IMAGE	SMF (%)	DICE COEFF (%)	RMSE	PSNR (dB)
Img_0	97.39	98.68	0.149	64.61
Img_1	93.82	96.74	0.209	61.69
Img_2	97.33	98.69	0.119	66.59
Img_3	95.01	97.44	0.071	71.09
Img_4	99.60	99.85	0.024	80.38
Img_5	98.37	98.83	0.035	77.26

Img_6	94.85	97.21	0.087	69.32
Img_7	98.94	99.57	0.084	69.63
Img_8	97.42	98.57	0.176	63.22
Img_9	96.47	98.21	0.084	69.63

Table II .Tumor Detection Accuracy

IMAGE	SMF (%)	DICE COEFF (%)	RMSE	PSNR (dB)
Img_0	95.65	96.97	0.0103	86.84
Img_1	97.79	97.91	0.0161	83.99
Img_2	98.43	97.95	0.0163	83.94
Img_3	99.06	99.51	0.0117	87.75
Img_4	96.74	97.73	0.0191	82.49
Img_5	95.12	97.17	0.0331	77.72
Img_6	98.65	99.24	0.0310	78.30
Img_7	91.91	95.67	0.0520	72.81
Img_8	90.12	91.47	0.0599	72.58
Img_9	97.52	98.78	0.0160	84.00

The tumor burden analysis performed on the output of proposed algorithm and reference image of the GROUP 1 images are recorded as shown in Table III. According to this table, tumor burden values less than or equal to 5, then the tumor correspond to benign. While tumor burden value greater than 5, then the tumor is malignant. Severity of cancer is calculated using tumor burden analysis, where malignant tumor is found to be highly cancerous and benign tumor is found to be less cancerous. The tumor burden analysis performed on GROUP 2 images are recorded in Table IV. From the table, it is clear that tumor can be correctly classified as either malignant or benign on the basis of tumor burden analysis.

Table III .Tumor Burden (TB) Analysis of Group 1

IMAGE	TB FROM ALGORITHM (%)	TB FROM REFERNCE IMAGE (%)	CATEGORY
Img_0	2.83	2.87	Benign
Img_1	2.34	2.36	Benign
Img_2	4.67	4.84	Benign
Img_3	2.77	2.81	Benign
Img_4	1.97	2.01	Benign
Img_5	3.04	3.11	Benign
Img_6	4.99	5.00	Benign
Img_7	2.02	2.24	Benign
Img_8	1.33	1.57	Benign
Img_9	3.41	3.52	Benign
Img_10	12.06	12.45	Malignant
Img_11	24.78	24.78	Malignant
Img_12	14.33	15.08	Malignant
Img_13	9.15	9.99	Malignant
Img_14	10.79	10.80	Malignant
Img_15	16.64	16.79	Malignant
Img_16	7.62	7.84	Malignant
Img_17	10.31	11.91	Malignant
Img_18	26.95	28.85	Malignant
Img_19	18.58	18.83	Malignant

Table IV. Tumor Burden (TB) Analysis of Group 2

IMAGE	TB (%)	CATEGORY
Img_0	8.46	Malignant
Img_1	18.24	Malignant
Img_2	4.99	Benign
Img_3	9.80	malignant
Img_4	2.04	Benign
Img_5	2.01	Benign
Img_6	3.32	Benign
Img_7	16.64	Malignant
Img_8	4.23	Benign
Img_9	15.01	Malignant

V. CONCLUSION

This paper provides an automated method to segment liver portion, followed by tumor area from abdominal CT image and a method for classification of benign and malignant tumor by using Support Vector machine. This method is tested on 30 different abdomen CT images. The pre-processing is done by using median filter to remove noise from the image. Then the liver is segmented by using a localized region based active contouring algorithm and the region growing method is used for the detection of tumor. Then the tumor is classified as benign or malignant by using SVM approach. The reference image is created by using ECOGNITION SOFTWARE. The experimental results show that liver segmentation and tumor detection is accurate. The SVM leads to provide better accuracy.

REFERENCES

[1] L. Soler,, H. Delingette, G. Malandain, J. Motagnat, N. Ayache, C. Koehl, O. Dourtheb, B. Malassagne, M. Smith, D. Mutter, J. Marescaux: "Fully automatic anatomical, pathological, and functional segmentation from ct". Proc. SPIE Medical Imaging 3979 (2000) 246–255.

[2] H. Lamecker, T. Lange, M. Seebass: "Segmentation of the liver using a 3d statistical shape model". Technical Report ZIB-Report 04-09, Konrad-Zuse-Zentrum fr Informationstechnik Berlin (April 2004).

[3] P. Campadelli, E. Casiraghi, G. Lombardi, "Automatic Liver Segmentation from Abdominal CT Scans", 14th International Conference on Image Analysis and Processing-ICIAP 2007, IEEE Computer Society, pp. 731- 736, 2007.

[4] T. Okada, R. Shimada, Y. Sato, M. Hori, K. Yokota, M. Nakamoto, Y.W. Chen, H. Nakamura, "Automated Segmentation of the Liver from 3D CT Images Using Probabilistic Atlas and Multi-level Statistical Shape Model," MICCAI 2007, LNCS, vol.4791, pp. 86-93, 2007

[5] J A Sethian,.: "Level Set Methods and Fast Marching Methods". Cambridge University Press (1999).

[6] T. E. Chan, B. Y. Sandberg, and L. A. Vese, "Active contours without edges for vector-valued images," *J. Visual Commun. Image Represent.*, vol. 11, pp. 130–141, Jun. 2000. [7] Pohle, R., Toennies, K.D.: Segmentation of medical images using adaptive region growing. Proc. SPIE Medical Imaging 4322 (2001) 1337–1346.

[8] K.-S. Seo, "Automatic Hepatic Tumor Segmentation Using Composite Hypotheses," in *Image Analysis and Recognition*, vol. 3656, 2005, pp.922-929.

[9] K. K. Dipak, B. Saptarshi, M. Sreeja, M. Sampita, S. Atreyee, S. Souptik, M. Dibya and ArunaChakraborty, "Automatic Detection and Size Measurement of Hepatic Lesions", International Journal of Wisdom Based Computing, Vol. 1 (3), December 2011.

[10] H. Nader, M. Mohiy, M. Khalid, "Fully automatic liver tumor segmentation from abdominal CT scans" IEEE 2010.

[11] A.G Sajith, S. Hariharan, "Medical Image Segmentation Using CT

Scans-A Level Set Approach",International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-2, Issue-6, May 2013.

[12] S. Lankton and A. Tannenbaum, "Localizing region-based active contours," *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2029–2039, Nov. 2008.

[13] G. Megha, J. P. Gawande," Automated Segmentation of Liver and Tumourand Feature Extraction from Abdominal Ct Images Using Region Growing Method" *Proceedings of 10th Irf International Conference, 01st June-2014, Pune, India,*

[14] M V Sudhamani, G T Raju," Segmentation and Classification of Tumor in Computed Tomography Liver Images for Detection, Analysis and Preoperative Planning", *International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-4 Number-1 Issue-14 March-2014.*

Priya V. received the B. Tech Degree in Computer Science and Engineering from Cusat University in 2012. She is currently pursuing M. Tech Degree in Computer and Information Science at College of Engineering Munnar. Her current area of interest is in image processing.

Biju V. G. received his B. Tech Degree in Electronic Engineering from Cochin University of Science and Technology, Kerala. M. Tech Degree in Electronics and Communication (Optoelectronics and Optical Communication) from University of Kerala. He is an Associate Professor in the Department of Electronics and Communication at College of Engineering Munnar, Idukki, Kerala. His areas of interest in research include image processing, Signal Proc essing, Pattern recognition, and Bioinformatics.