ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Enhanced Detection of Lung Cancer using Hybrid Method of Image Segmentation

L Uma Maheshwari
Department of ECE,
Stanley College of Engineering and Technology for Women, Hyderabad - 500001, India.

Udayini Dikkala
Department of ECE,
Stanley College of Engineering and Technology for Women, Hyderabad - 500001, India.

Abstract:
Effective identification of lung cancer at an initial stage is an important and crucial aspect for effective treatment and increase in survival rate. Lung cancer can be diagnosed with the help of MRI, CT scan, PET or X-ray. But CT scan images help in diagnosing the cancer at an early stage and include more detail than conventional X-ray images. CT scan images from database LIDC-TCIA are considered for research purpose. Several segmentation methods have been used to detect lung cancer at an early stage. Thresholding and Marker Controlled Watershed algorithm provides lower accuracy. To improve accuracy and determine the stage of cancer, hybrid method of image segmentation has been adopted. The system comprises of preprocessing, segmentation, feature extraction and classification steps. Classification is done using Support Vector Machine (SVM) to determine the severity of disease, i.e., in which stage the cancer is present. MATLAB R2010a has been used for simulation to detect lung cancer from CT scan images.

Keywords:
Fuzzy C-Means (FCM), Lung Cancer, Support Vector Machine (SVM), Watershed Segmentation.

I. INTRODUCTION:

Currently the image processing mechanisms are used widely in different medical areas for providing earlier detection and treatment stages. The early detection of the disease increases the chances of an effective treatment. In 2005, approximately 1,372,910 new cancer cases were predictable and about 570,280 cancer deaths were expected to occur.

It was expected that there would be 163,510 deaths from lung cancer, which forms 29% of all cancer deaths [1]. Several low pre-processing techniques are used to segment lung and other body organs [2-6]. When cells start to grow out of control in abnormal proportions, cancer begins in a part of the body [3]. Lung cancer is a disease of abnormal cells multiplication and development into a tumour. Cancer cells continue to increase and form new, abnormal cells [4]. Many of the patients notice their disease when it is too late and the surgery is not simply possible. Diagnosis is mainly based on CT scan images [5]. Cancerous tumour starts in a certain part of lung and is called primary lung cancer [6]. Following are the types of this lung cancer and these are divided into two main types:

1. Small cell cancer
2. Non-small cell cancer

The current work focuses on finding the tumour and its stage of development. In this Marker-controlled Watershed segmentation is used to isolate a lung of a CT image.

II. PROPOSED SYSTEM:

In this system, available lung CT scan images are passed through the system which is having following stages: pre-processing stage, segmentation stage, feature Extraction stage and classification. The block diagram of the flow of the process is shown in Figure 1 below.

Cite this article as: L Uma Maheshwari & Udayini Dikkala, "Enhanced Detection of Lung Cancer using Hybrid Method of Image Segmentation", International Journal & Magazine of Engineering, Technology, Management and Research, Volume 5, Issue 1, 2018, Page 7-13.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

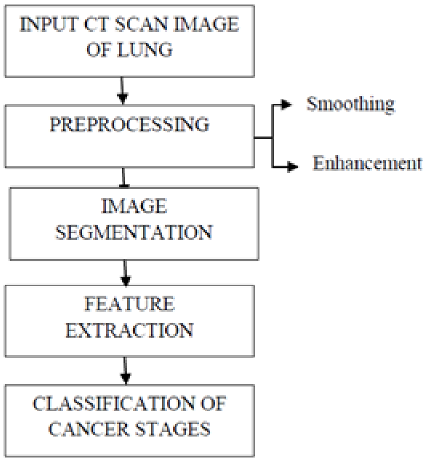
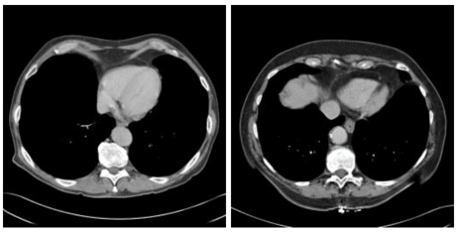


Figure 1. Block diagram of Hybrid method of segmentation

The Gaussian filter is used to smooth the input image in the preprocessing stage. Further, for image enhancement, in the pre-processing stage, Gabor filter is used. The process of Fuzzy C Means, followed by Marker-Controlled Watershed transform is used for the segmentation purpose. After image segmentation, the GLCM features such as perimeter, area and eccentricity are extracted from the detected tumour. Binarization process is done to decide whether it is cancerous tumour or not. Also, if there is cancerous tumour, the cancer stage is identified using Support Vector Machine.

A. Input CT scan images of lung:

The CT scan images which are used for processing are collected from the hospitals. This image dataset contains lung CT scan images with tumour and without tumour. Figure 2 shows some of the lung CT scan images (a) without tumour and (b) with tumour.



(a) (b)

Figure 2. CT scan image (a) Lung without tumour (b)Lung with tumour

B. Preprocessing:

For smoothing, Gaussian filter is applied on the input image. Gaussian smoothing is very effective for removing noise or components of high frequency from the image. The result is a smoothed image as shown in Figure 3. As seen, Gaussian filter is a low pass filter. Smoothing reduces the noise and results in a more accurate intensity surface. Next part in pre-processing is image enhancement stage. The function of image enhancement stage is to highlight important information of the image. In this step, better visual effects are performed on the image which enhance the human eyes' distinguish ability of information. It is a way to improve the class of image, so that the final output image is better than the original one. For image enhancement, Gabor filter is used and the result is as shown in Figure 4. The Gabor function is a very helpful tool in image processing, texture analysis. It is a linear filter and its impulse response is derived from the multiplication of harmonic function and Gaussian function. It is a band pass filter. It is used to increase the contrast between the nodule areas and other structures around it.

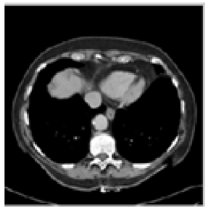


Figure 3. Smoothed Image

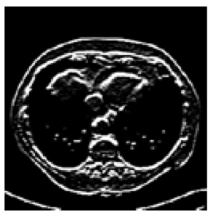


Figure 4. Enhanced Image using Gabor Filter

C. Image segmentation:

It is used to divide an image into different small regions or objects. It has many applications in the medical field for the segmentation of the 2D medical images.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

It is an important process for most image analysis following techniques. There are various methods available for image segmentation. In this paper, Fuzzy C-Means (FCM) and Marker Controlled Watershed segmentation methods are used. It is a technique wherein each data point belongs to a cluster to some degree that is specified by a membership grade. Generally, FCM algorithm proceeds by iterating the two indispensable conditions until a solution is reached. Each data point will be joined with a membership value for each class after FCM clustering. The objective of FCM is to determine the cluster centers and to produce the class membership matrix. In other words, it assigns a class membership to a data point, depending on the similarity of the data point to a scrupulous class relative to all other classes. It is based on minimization of the following objective function

$$J_m = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{m} \|x_i - c_j\|^2 \qquad , 1 \leq m < \infty$$

where N is number of clusters and C is number of centroids m is any real number greater than 1, uij is the degree of membership of xi in the cluster j, xi is the ith of d-dimensional measured data, cj is the d-dimension center of the cluster, and ||*|| is any norm expressing the similarity between any measured data and the center. Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership uij and the cluster centers cj. The output of FCM is as shown in Figure 5.

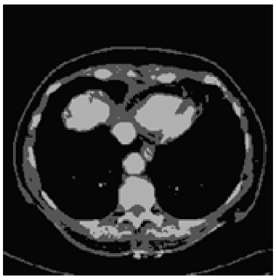


Figure 5. FCM based segmented Image

The concept of Watershed is well known in topography. Watershed segmentation is used to extract the region minimum value from an image. It determines the corresponding to the dividing line with the least value. Dividing line in the image gives the rapid change of boundary. This transform finds catchment basins and watershed edge lines in the image. It treats the image as a plane, where light pixels are high and dark pixels are low. The important drawback associated to the Watershed Transform is the over segmentation that usually results. The output of watershed segmented image is shown in Figure 6.

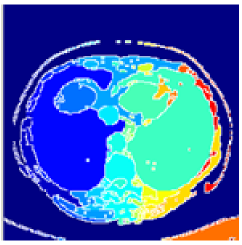


Figure 6. Watershed segmented Image

To overcome the drawbacks of this Watershed Segmentation i.e. over segmentation, the marker based watershed segmentation technique is used. It can segment boundaries from an image. Morphological operations are performed on the watershed segmented image to get final segmented image. Here the method is to use morphological operations called opening by reconstruction and closing by reconstruction to clean up the image. These operations will generate flat maxima inside each object which is located using image regional maxima. The output of Marker based watershed segmented image is shown in Figure 7.

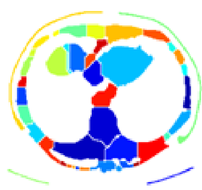


Figure 7. Marker-Controlled Watershed based segmented Image

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

After segmentation process, binarization process is done. In this approach, the total number of black pixels & white pixels are counted. If the total number of black pixels of input image is more than threshold, then the tumour is normal tumour. Otherwise, if the total number of the black pixels is less than the threshold, then the tumour is a cancerous tumour. By using this result, sensitivity and accuracy are calculated. To find these parameters, first calculate some of the conditions like true positive, false negative, true negative and false positive.

Sensitivity = TP/ (TP + FN)

Accuracy = (TN + TP)/ (TN + TP + FN + FP)

Here, TP is True Positive, TN is True Negative, FN is False Negative and FP is False Positive. Sensitivity is defined as amount of true positives that are correctly recognized by a diagnostic test. Accuracy is defined as the amount of true results, which is either true positive or true negative. It measures the degree of reliability of a diagnostic test on a condition.

D. Feature Extraction and Classification:

The GLCM features, for example, area, perimeter, eccentricity PSNR, MSE, homogeneity, entropy, energy, contrast and correlation are extracted and according to these features, tumour will be classified into the stages of cancer using Support Vector Machine. Support Vector Machines are supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification. The basic SVM takes a set of input data and for each given input, predicts which of two classes forms the input, making it a non-probabilistic binary linear classifier. SVM uses a kernel function which maps the given data into a different space; the separations can be made even with very complex boundaries. The different types of kernel function include polynomial, RBF, quadratic, Multi-Layer Perceptron (MLP). Staging involves evaluation of a cancer size and its penetration into surrounding tissues as well as presence or absence of metastasis in the lymph nodes or other organs [7]. Stages from I to IV in order of severity:

- Stage I: Cancer is confined to the lung
- Stage II and III: Cancer is confined to the chest
- Stage IV: Cancer has spread from the chest to other parts of the body.

By analysis, it is found that the performance of the hybrid method, i.e., FCM along with Marker Controlled Watershed Transform, is much superior than the existing algorithms. The results of the analysis have proved 81% of accuracy in hybrid method algorithm by detecting the tumours from the CT scan lung images when compared to the earlier algorithms, with an accuracy of upto 74%.

III. RESULTS:

Figure 8 shows all the outputs at each of the intermediate steps in the hybrid segmentation method. The original image and the segmentation outputs of CT scan image with vascularised tumour are as shown in Figure 8(a).

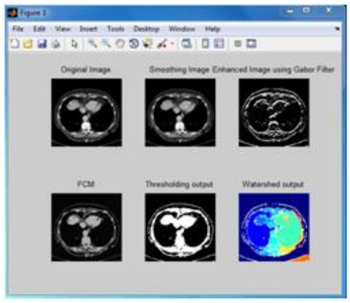


Figure 8(a)

The Marker-Controlled Watershed output is shown in Figure 8(b).

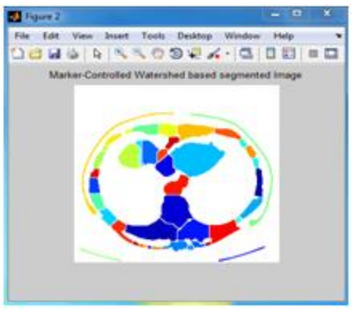


Figure 8(b) Marker Controlled Watershed result

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Manual selection of ROI on the binarised image is shown in Figure 8(c). Once segmentation is done, the type of tumor (cancerous or normal) is analyzed using binarization. If it is cancerous, the system asks the user to select ROI as in Figure 8(c) to determine the features of the tumor. Features like area, perimeter, eccentricity, homogeneity, entropy, PSNR are determined. The selected ROI is displayed as separated lung nodule as shown in Figure 8(d).

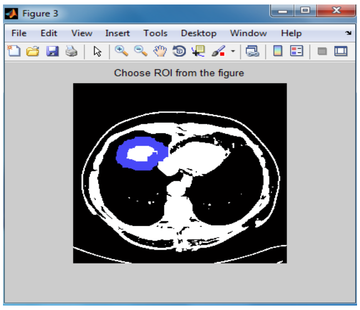


Figure 8(c) ROI for area in focus

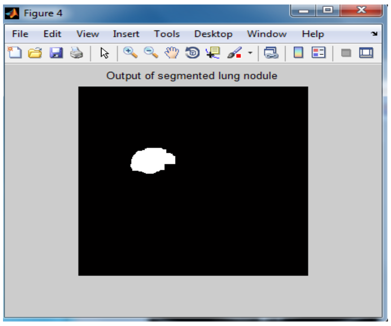


Figure 8(d) Separated lung Nodule

Feature extraction is done based on the parameters derived for segmented lung nodule and classification is done using SVM classifier, the result of which is shown in Figure 8(e).

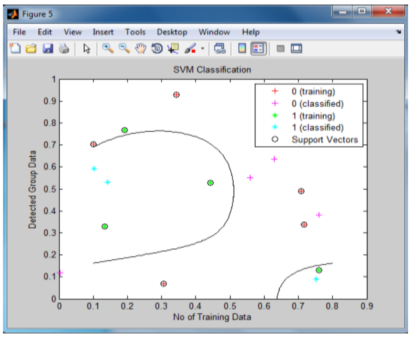


Figure 8(e) SVM Classifier result

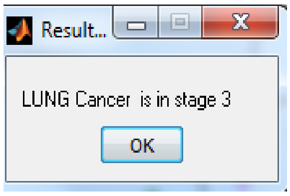


Figure 8(f) Display of Cancer Stage

Result box displaying the stage of cancer is shown in Figure 8(f). The resulting Metrics is shown below in the Table 1 & 2.

Table 1: Metrics based on Confusion Matrix

Evaluation Metrics	Threshold based Segmentation	MCWT	FCM+MCWT
Sensitivity	0.57	0.71	0.79
Accuracy	0.57	0.74	0.81

Table 2: Metrics from the classifier

Parameters	MCWT	Hybrid method
Area (sq.mm)	921	1528
Perimeter (mm)	69.9120	72.7660
Eccentricity	0.9895	0.9955
Entropy	0.0819	0.1197
Contrast	0.0402	0.0431
Energy	0.9709	0.9705
Homogeneity	0.9988	0.9988
MSE	0.0186	0.0081
PSNR	10.8918	18.0185

It can be seen that area, perimeter, eccentricity, entropy and contrast are increased in hybrid method of segmentation. Energy dissipated had decreased. MSE and PSNR define the quality. There is an increase in PSNR and decrease in MSE when compared with Marker Controlled Watershed Algorithm.

IV. CONCLUSION AND FUTURE SCOPE:

The system consists of pre-processing, segmentation, feature extraction and final classification. The proposed Marker Controlled Watershed segmentation technique separates the objects in contact in the image. FCM is one of an accepted clustering method and has been broadly applied for medical image segmentation. It gives higher accuracy compared to the thresholding algorithm and Marker Controlled Watershed Algorithm. So it is an efficient method of segmentation. From the extracted region of interest, three main features are extracted i.e., area, perimeter and eccentricity.

These three features help to identify the stage of lung cancer. The results indicate that the tumors are of different dimensions. By measuring the dimensions of the tumor the lung cancer stage can be detected accurately using the proposed method. Also for classification purpose, Support Vector Machines are an attractive approach to data modeling. Present technique gives very promising results in comparison to other used techniques even though it requires high computational time. The preprocessing step, where filters are used, can be improved with the use of Discrete Wavelets and Curvelets.

REFERENCES:

[1] Sayali Satish Kanitkar, N. D. Thombare, S. S. Lokhande, "Detection of lung cancer using marker controlled watershed transform", International Conference on Pervasive Computing (ICPC), 2015.

[2] Nihad Mesanovic, Haris Huseinagic, Matija Males, Mislav Grgic, Emir Skejic, Muamer Smajlovic, "Automatic CT Image Segmentation of the Lungs with Region Growing Algorithm".

[3] Sayani Nandy, Nikita Pandey, "A Novel Approach of Cancerous Cells Detection from Lungs CT Scan Images", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 2, Issue 8, August 2012.

[4] Prof. Samir Kumar Bandyopadhyay, "Edge Detection from CT Images of Lung", International Journal of Engineering, Science & Advanced Technology, 2012, Volume-2, Issue-1, pg 34–37.

[5] Fatm Taher, Naoufel Werghi and Hussain Al-Ahmad, "Extraction of Sputum Cells using Thresholding Techniques for Lung Cancer Detection", International Conference on Innovations in Information Technology (IIT), 2012.

[6] K. Parvati, B. S. Prakasa Rao and M. Mariya Das, "Image Segmentation Using Gray-Scale Morphology and Marker-Controlled Watershed Transformation", Hindawi Publishing Corporation Discrete Dynamics in Nature and Society Volume, 2008, Article ID 384346.

Author Details:

L. Uma Maheshwari

Received the B.E degree in Electronics and Communication Engineering from Osmania University, India. She is pursuing Masters in Embedded Systems from Stanley College of Engineering and Technology for Women, Hyderabad, Telangana, India.

Udayini Dikkala

Received the B.E. degree in Electronics and Communication Engineering from Andhra University, India and the M.S. degree in Electrical Engineering from the University of Houston, Texas, USA, in 1994 and 1996 respectively. She worked in the software industry for a period of more than 4 years on various platforms as a System Analyst.

Since 2011, she has been with Stanley College of Engineering and Technology for Women, Hyderabad, India, where she is currently working as Assistant Professor in the Department of Electronics and Communication Engineering. Her research interests include Medical Image Processing, Signal Processing and Digital Electronics. She is a member of IEEE, IETE and ISTE.