iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** VII **Month of publication:** July 2015

DOI:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

MRI Classification and Segmentation of Cervical Cancer to Find the Area of Tumor

S. S. Dhumal[1], S. S. Agrawal[2]
[1]Department of EnTC, University of Pune, India

Abstract— Automated classification and detection of tumors in different medical images is motivated by the necessity of high accuracy when dealing with a human life. The detection of the Cervical Tumor is a challenging problem, due to the structure of the Tumor cells. This proposed system presents a segmentation method spatial fuzzy clustering algorithm, for segmenting Magnetic Resonance images to detect the Cervical Tumor in its early stages and to analyze anatomical structures. The vector machine will be used to classify the whether test image of Cervical MRI is normal or abnormal. Here Dual Tree CWT multi scale decomposition is used to analysis texture of an image. The segmentation results will be used for early detection of Cervical Tumor which will improve the chances of survival for the patient. To implement an automated Cervical Tumor classification decision making was performed in two stages: feature extraction using GLCM and the classification using SVM. The performance of this classifier was evaluated in terms of training performance and classification accuracies. The simulated results will be shown that classifier and segmentation algorithm provides better accuracy than previous method.
Keywords— Cervix, classification, feature extraction, magnetic resonance imaging (MRI), segmentation.

I. INTRODUCTION

Cervical cancer accounts for approximately 500,000 new cases and 274,000 deaths every year worldwide [1]. It is the second leading cause of death from cancer among women. However significant progress in reducing these deaths have occurred in the last few decades and it can be attributed to increasingly accurate early screening test. The screening tests consist of microscopic examination of exfoliated cells from the transformation zone of the cervix. The screening tests still need to be made easy and prevalent to prevent cancer deaths especially in the developing world. Around 80% of all cervical cancer deaths occur in the developing world [2]. In this regard a medical image processing approach towards automatic detection of the presence and level of cancer in test samples becomes highly desirable.

Tumor is an uncontrolled growth of tissues in any part of the body. Tumors are of different types and they have different Characteristics and different treatment. Most Research in developed countries show that the number of people who have cervical tumors died due to the fact of inaccurate detection.

The objectives of this methodology are listed below.

To decompose the test and training MR image using Dual Tree Complex Wavelet Transform

To extract the features of the compressed image using Grey Level Co-occurrence Matrix.

To classify the cervix image using Support Vector Machine approach.

To segment the abnormal image using Spatial Fuzzy Clustering.

This chapter introduces the main concept of what cervical cancer is and how it is dangerous if not cured in time. It even explains us with advanced automation; data is evaluated in real time to extract the information, which is further used for medical application.

II. METHODOLOGY

MRI Cervical image Classification and anatomical structure analysis is carried out by using the following methodologies step by step in given sequence.

MRI of Cervix

Dual Tree CWT multi scale decomposition

GLCM Feature Extraction

SVM Training and Classification

Spatial Fuzzy Clustering Method

Morphological processing

Performance analysis

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

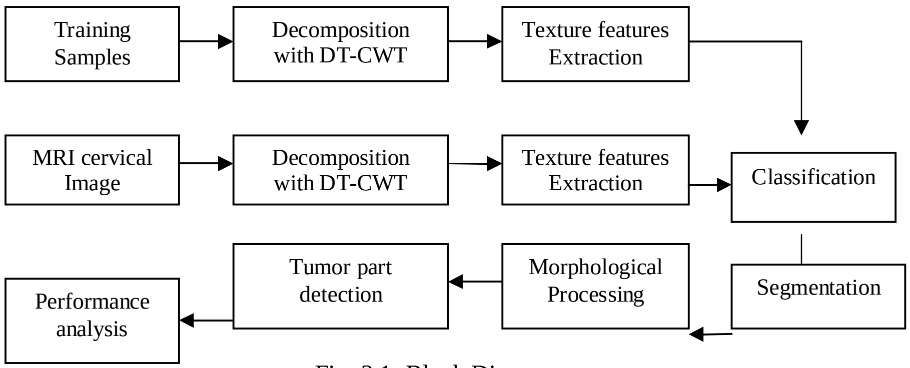

Fig. 3.1: Block Diagram

Fig. 3.1 represents the block diagram of the methodology used for the project. All the methodologies used in this project are explained in detail below.

A. MRI (Magnetic Resonance Image)

MR images are taken as test images. Magnetic resonance imaging (MRI) is high-quality medical imaging, particularly for cervical imaging. MRI inside the human body is helpful to see the level of detail. Doctors have major technical and economic importance of reliable and fast detection and classification of cervical cancer, based on common practices. Most of the technicians are slow, less responsible, and that's hard to quantify possess a degree of subjectivity. For the early detection of cervical tumors there are many imaging methods for diagnostics purpose are presented.

B. DT-CWT (Dual Tree Complex Wavelet Transform)

DT-CWT technique is used for decomposition of the input MR image. The dual-tree complex wavelet transform (CWT) is a relatively recent enhancement to the discrete wavelet transform (DWT), with important additional properties: It is nearly shift invariant and directionally selective in two and higher dimensions. It achieves this with a redundancy factor of only 2d for d-dimensional signals, which is substantially lower than the undecimated DWT. The multidimensional (M-D) dual-tree CWT is non-separable but is based on a computationally efficient, separable filter bank (FB). This tutorial discusses the theory behind the dual-tree transform, shows how complex wavelets with good properties can be designed, and illustrates a range of applications in signal and image processing. We use the complex number symbol C in CWT to avoid confusion with the often-used acronym CWT for the (different) continuous wavelet transform.

C. GLCM (Grey Level Co-occurance Matrix)

Features or the MR images of cervix are obtained by using the technique of GLCM. A gray level co-occurrence matrix (GLCM) can be constructed for any single channel image [3]. The GLCM is a square matrix where the number of rows and columns equals the number of gray levels in the original image. The four features contrast, correlation, energy, and homogeneity were used in to extract the features of the cervix MRI.

First, **contrast** K is defined as

$$\mathbf{K} = \sum_{i,j} \binom{n}{k} (i-j)^2 p(i,j) \qquad (2.1)$$

This texture feature measures the contrast in gray level from one pixel to its neighbour. The contrast ranges from zero for a constant image to $(G-1)^2$, where G is the number of gray levels.

Second, **correlation** R is defined by,

$$\mathbf{R} = \sum \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma i \sigma j} \qquad (2.2)$$

R measures the correlation between intensities in neighbouring pixels. μi is the row average and μj is the column average of the GLCM. σi and σj are the standard deviations of row i and column j in the GLCM, respectively. Third, **energy** E, also known as the angular second moment, is simply the sum of squares of all the elements in the GLCM,

$$\mathbf{E} = \sum_{i,j} p(i,j)^2 \qquad (2.3)$$

Energy measures uniformity, and ranges from zero to unity for a uniform image.

Fourth, **homogeneity** H is defined as,

Volume 3 Issue VII, July 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

$$\mathbf{H} = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \qquad (2.4)$$

H is a measure of the closeness of the elements in the GLCM to the diagonal. It ranges from zero to one. Homogeneity is unity for a diagonal GLCM, that is, if all pixels in the original image have the same value as their neighbour.

D. SVM (Support Vector Machine)

The main objective of this methodology is classification of the MR images of cervix with the training sample images to verify if there is any anatomical structure present in given image. This is done using an approach SVM. Support Vector Machine (SVM) is a classification and regression prediction tool that uses machine learning theory to maximize predictive accuracy while automatically avoiding over-fit to the data. Support Vector machines can be defined as systems which use hypothesis space of a linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory.

The general framework to measure the accuracy of a SVM on a given database is composed of the following stages:

Pre-processing of the images in the database
Separation of the database in training and test sets
Choice of the representation of the input data
Choice of the way of training, which includes:
Method of multi-class training
Value of the penalty term C
Choice of the kernel
Training
Test and evaluation of the performance.

Classification in SVM is an example of Supervised Learning. Known labels help indicate whether the system is performing in a right way or not. This information points to a desired response, validating the accuracy of the system, or be used to help the system learn to act correctly. A step in SVM classification involves identification as which are intimately connected to the known classes. This is called feature selection or feature extraction. Feature selection and SVM classification together have a use even when prediction of unknown samples is not necessary. They can be used to identify key sets which are involved in whatever processes distinguish the classes [4]. If data is linear, a separating hyper plane may be used to divide the data. However it is often the case that the data is far from linear and the datasets are inseparable. To allow for this kernels are used to non-linearly map the input data to a high-dimensional space. The new mapping is then linearly separable. A very simple illustration of this is shown below in fig. 2.2.

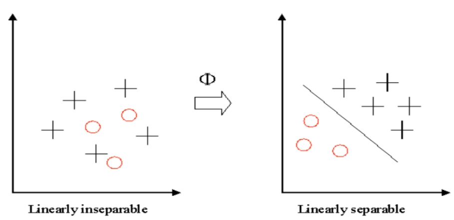

Fig. 2.2: SVM classification

Accuracy, sensitivity, specificity, and fraction of support vectors were used to evaluate the performance of the SVM models. Accuracy, sensitivity, and specificity are defined as

$$\text{Accuracy} = (TP+TN)/(TP+TN+FN+FP) \qquad (2.5)$$

$$\text{Sensitivity} = TP/(TP+FN) \qquad (2.6)$$

$$\text{Specificity} = TN/(TN+FP) \qquad (2.7)$$

Where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives, and FN is the number of false negatives. Using sensitivity / specificity in addition to accuracy ensured that both patient groups were well classified [5]. SVM models with few support vectors were preferentially selected to reduce the risk of over fitting.

E. Spatial Fuzzy Clustering

The technique used in this paper for the segmentation of the classified image is spatial fuzzy clustering technique. If the image is classified as normal by the SVM classification then there is no need of further segmentation. But if the image is classified as abnormal then we further need to segment the image. As the image is classified as abnormal, it means that it consists of some abnormal tissue growth in particular area which is formation of tumor. Thus, we need to segment the image so as to get the cluster formation view to detect the tumor area accurately. The fuzzy logic is a way to processing the data by giving the partial membership value to each pixel in the image. The membership value of the fuzzy set is ranges from 0 to 1. Fuzzy clustering is basically a multi valued logic that allows intermediate values i.e., member of one fuzzy set can also be member of other fuzzy sets in the same image. There is no abrupt transition between full membership and non membership. The membership function defines the fuzziness of an image and also to define the information contained in the image. These are three main basic features involved in characterized by membership function. They are support, Boundary. The core is a fully member of the fuzzy set. The support is non membership value of the set and boundary is the intermediate or partial membership with value between 0 and 1 [6].

F. Morphological Processing

According to the need of the next level this step converts the image. It performs filtering of noise and other artifacts in the image and sharpening the edges in the image. RGB to grey conversion and Reshaping also takes place here. It includes median filter for noise removal. The possibilities of arrival of noise in modern MRI scan are very less. It may arrive due to the thermal effect. The main aim of this paper is to detect and segment the tumor cells. But for the complete system it needs the process of noise removal. For better understanding the function of median filter, we added the salt and pepper noise artificially and removing it using median filter. Morphology is the study of shapes and structures from a scientific perspective. Morphological filters are formed from the basic morphology operations. A structuring element is mainly required for any morphological operation. Morphological operations operate on two images, structuring element and the input image. Structuring elements are small images that are used to probe an input image for properties of interest. Origin of a structuring element is defined by the centre pixel of the structuring element. In morphology, the structuring element defined will pass over a section of the input image where this section is defined by the neighbourhood window of the structuring element and the structuring element either fits or not fits the input image.

G. Performance Analysis

This step helps to find out the area of tumor present in the abnormal image of cervix. This area is found out in mm^2. In this process the tumor area is calculated using the binarization method. That is the image having only two values either black or white (0 or 1). Here 256x256 jpeg image is a maximum image size. The binary image can be represented as a summation of total number of white and black pixels [6].

III. IMPLEMENTATION

The implementation chapter gives all the algorithms and formulas which can be used to find out the correct accuracy of the system.

A. Spatial Fuzzy

First stem of implementation is the clustering algorithm which allows one piece of data may be member of more than one clusters. It is based on reducing the following function

$$\mathbf{Y_m} = \sum_{i=1}^{N} \sum_{j=1}^{C} \mathbf{M_{ij}^{m}} \, \|\mathbf{xi} - \mathbf{cj}\|\text{^}\mathbf{2} \qquad (3.1)$$

Where

M : any real number greater than 1,

Mij : degree of membership of X; in the cluster j,

x, : data measured in d-dimensional,

Rj : d-dimension center of the cluster,

The update of membership Mij and the cluster centers R, are given by:

$$\mathbf{M_{ij}} = \frac{\mathbf{1}}{\sum_{K=1}^{C} \left(\frac{\|\mathbf{xi}-\mathbf{cj}\|}{\|\mathbf{xi}-\mathbf{cj}\|}\right)\text{^}\frac{\mathbf{2}}{\mathbf{m-1}}} \qquad (3.2)$$

$$\mathbf{R_j} = \frac{\sum_{i-1}^{N} \mathbf{xi.M_{ij}^{m}}}{\sum_{i=1}^{N} \mathbf{M_{ij}^{m}}} \qquad (3.3)$$

The above process ends when,

$$\max_{ij}\left\{\left|M_{ij}^{(K+1)} - M_{ij}^{k}\right|\right\} < \delta \qquad (3.4)$$

Where

δ = termination value or constant between 0 and 1,

K= no of iteration steps.

The algorithm for this spatial fuzzy clustering contain following steps:

Step 1. Initialize M = $[M_{ij}]$matrix, $M^{(0)}$

Step 2. At k-step: calculate the centers vectors $R^{(k)} = [R_j]$ with $M^{(k)}$

$$R_j = \frac{\sum_{i=1}^{N} xi\,.M_{ij}^{m}}{\sum_{i=1}^{N} M_{ij}^{m}} \qquad (3.5)$$

Step 3. Update $U^{(k)}$, $U^{(k+l)}$

$$M_{ij} = \frac{1}{\sum_{K=1}^{C}\left(\frac{||xi-cj||}{||xi-cj||}\right)^\wedge\frac{2}{m-1}} \qquad (3.6)$$

Step 4. If $||M^{(k+1)} - M^{(k)}|| < \xi$ then STOP; otherwise return to step 2.

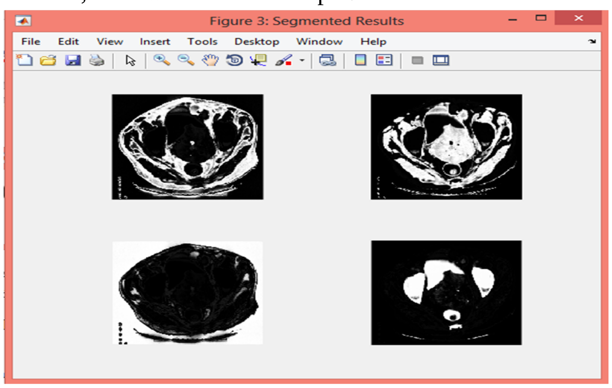

Fig. 3.1 Segmentation clusters

Fig. 3.1 is the result of segmentation by applying the algorithm given above it is mainly developed for the accurate prediction of tumor cells.

B. Tumor Area Detection

The tumor area can be found out by using the below explained steps along with its algorithm. The algorithmic steps involved for cervical tumor shape detection is as follows,

Start the process.

Get the MRI scan image input in JPEG format.

Check whether the input image is in required format and move to step 4 if not display error message.

If image is in RGB format covert it into gray scale else move to next step.

Find the edge of the grayscale image.

Calculate the number of white points in the image.

Calculate the size of the tumor using the formula (3.9).

Display the size and stage of tumor.

Stop the program.

In this process the tumor area is calculated using the binarization method. The binary image can be represented as a summation of total number of white and black pixels.

$$\text{Image, } \mathbf{I} = \sum_{W=0}^{255}\sum_{H=0}^{255}[\mathbf{f(0)} + \mathbf{f(1)}] \qquad (3.7)$$

Where,

Pixels = Width (W) X Height (H) = 256 X 256

f(0) = white pixel (digit 0)

f(l) = black pixel (digit 1)

$$\text{No. of White Pixels, } \mathbf{P} = \sum_{W=0}^{255}\sum_{H=0}^{255}[\mathbf{f(0)}] \qquad (3.8)$$

Where,

Volume 3 Issue VII, July 2015
ISSN: 2321-9653

P = number of white pixels (width*height)
1 Pixel = 0.264 mm
The area calculation formula is

$$\text{Size of tumor, } \mathbf{S} = [\sqrt{\boldsymbol{P}} * \mathbf{0.264}\ ]\ \text{mm}^2 \qquad (3.9)$$

P = no. of white pixels; W = width; H = height.
Thus, using the above algorithm we can find the area of tumor present in the MR image of cervix.

IV. RESULTS AND DISCUSSION

The results are obtained using software MATLAB. The version used for programming this project is MATLAB R2014b. A graphical user interface (GUI) is a user interface built with graphical objects, such as buttons, text fields, sliders, and menus. Thus, GUI is used as main parameter in programming. At the end area of tumor present in the cervix is measured in mm^2. Thus the area which was measured of the tumor was recorded as 15.3688 mm^2. The predicted tumor area is calculated at approximate reasoning step. Fig. 4.1 shows the output result for tumor area and its stage calculation. The stage of tumor is based on the area of tumor. We considered that, if the area is greater than 6 mm^2 it will be the critical position. The area calculated in our methodology is 15.3688 mm^2. Thus, we need to diagnose and take preventive steps as early as possible.

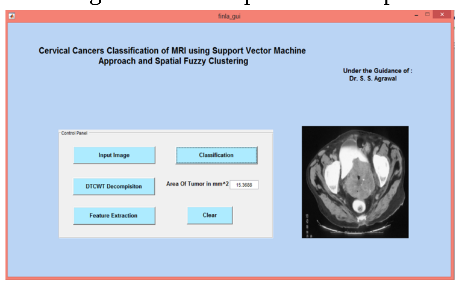

Fig. 4.1: Tumor Area Measurement

V. CONCLUSION

Pre-treatment MR Images of locally advanced cervical cancer contains information relevant for prediction of treatment response. The discriminating information was found in the spatial relations in tumor, given by GLCM texture features. For predicting outcome of chemo-radio therapy for cervical cancer patients, texture features from GLCMs derived from MRI appear to be better predictors. Spatial Fuzzy C Means method incorporates spatial information, and the membership weighting of each cluster is altered after the cluster distribution in the neighbourhood is considered. Finally approximate reasoning for calculating tumor area is done. The experimental results of the proposed method give more accurate result with accuracy of 74%, sensitivity of 76%, and specificity of 72%. Dynamic Enhanced MR Images (DCE-MRI) which provides insight into the vascular properties of tissue can be used instead of the MR Images. In future 3D assessment of cervix using 3D slicers with MATLAB can be developed.

VI. ACKNOWLEDGMENT

I am indeed thankful to my guide Prof. S. S. Agrawal for her able guidance and assistance to complete this paper; otherwise it would not have been accomplished. Author and co-author of this paper also would like to thank other researchers who provided the platform for this paper.

REFERENCES

[1] S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” IEEE Trans. on Neural Networks, vol. 4, pp. 570-578, July 1993.
[2] http://www.cervicalcanceraction.org/whynow/about.php, Accessed on 12th March 2013.
[3] R. M. Haralick, Shanmuga.K, and I. Dinstein, “Textural Features for Image Classification,” IEEE Trans. Syst. Man, Cybern. B, Cybern, vol. SMC3, no. 6, pp. 610-621, 1973.
[4] Nello Cristianini and John Shawe-Taylor, “An Introduction to Support Vector Machines and Other Kernel-based Learning Methods”, Cambridge University Press, 2000.
[5] D. L. Olson, and D. Delen, Advanced Data Mining Techniques, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008.
[6] J.selvakumar A.Lakshmi T.Arivoli, “Brain Tumor Segmentation and Its Area Calculation in Brain MR Images using K-Mean Clustering and Fuzzy C-Mean Algorithm,” IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012)pp.186-190 March 30, 31, 2012.