Research Paper

Volume : 4 | Issue : 5 | May 2015 • ISSN No 2277 - 8179

Classification of Mr Images of Cervical Cancer Using SVM and ANN

Engineering

KEYWORDS : Cervical cancer, Automatic support system, Thresholding, FCM,classification

Aabha S. Phatak — Dept. of Electronics and telecommunication Vishwakarma Institute of Technology, Pune, India.

Prof. Dr. Ashwini B. Barbadekar — Dept. of Electronics and telecommunication Vishwakarma Institute of Technology, Pune, India.

ABSTRACT

Cervical cancer can be considered to be one of the key causes of cancer affecting women worldwide and its rate is rising . Early diagnosis can prove to be really instrumental in decreasing the cervical cancer mortality. Diagnosis using medical image analysis is gaining fast recognition. Thus, an automated diagnostic system using image processing techniques has been proposed in the paper, which may prove to be an aid to the radiologists and play a key role in early detection of cancer . The paper presents segmentation and classification techniques using which the envisioned automated decision support system functions.

INTRODUCTION

Uterine cervical cancer is the second most common form of cancer affecting women worldwide. Cervical cancer tends to be gradually developing, unlike other more aggressive cancers, with a development rate of several years. When detected and treated early, cervical cancer has a recovery rate of nearly 100 percent. This has lead to an increasing demand for early detection methods with heightened diagnostic accuracy. Colposcopy, pap smear analysis, lbc or thinprep pap test, biopsy are some of the well established diagnostic methods which are currently in practice, to detect cancerous and precancerous tissue through either visual inspection of the cervix or cervical cells. Envisioned automated decision support system aim to aid the physician in their diagnosis by exploiting the image processing algorithms on the magnetic resonance images of cervical cancer, which are being used as input. This will result in more calculable diagnosis, reduction of the variability among cystoscopists and the minimization of dispensable biopsies. Furthermore, the existing techniques, even when carried out by experts, are subjective, not easy to replicate and a scenario cannot be ruled out where they can even be erroneous. Thus automated detection and classification of tumors in different medical images may prove to be instrumental in aiding the physicians and is clearly driven by the necessity of high accuracy when dealing with a human life .

LITERATURE SURVEY

A. Existing Methods/Approaches

The traditional/conventional methods which were/are followed , in brief are as under :

Dr. Georgios Papanikolaou , a greek doctor , first invented a method of cervical screening , used to detect potentially precancerous and cancerous processes in the endocervical canal (transformation zone) of the female reproductive system.The Papanicolaou test (abbreviated as Pap test, known earlier as Pap smear, or smear test) is performed by using a speculum used for opening the vaginal canal and then collecting cells from the outer opening of the cervix of the uterus and the endocervix. Examination of cells under a microscope is then carried out to look for abnormalities[1]

In an effort to improve the clarity and consistency of the pap test, a method was developed where the sample is collected from the patient in the same manner, but is instantly cleansed in a vial of preservative solution. The sample is sent to the lab, where its processing is done using an automated instrument that removes non-diagnostic debris and transfers a representative sample of the cells in a thin layer on the slide .The ThinPrep Pap Test thus resulted in an improved cellular definition and allows cytotechnologists to better use their classification skills instead of spending time locating cells of interest[2]

FocalPoint with Location Guided Screening (Focal Point GS) was proposed, which used Location Guided Screening to automate the microscopic analysis of slides . The Focal Point instrument is interfaced to a Slide Wizard platform and networked to microscopes equipped with computer-controlled automated stages for fast relocation of fields of interest on microscopic slides. During the initial screening process, and for each slide screened, Focal Point GS identifies and stores 15 fields of interest in which it has calculated a higher probability of abnormality. The FocalPoint GS communicates the location coordinates to the computer-controlled microscope stage via the Slide Wizard platform. The fields are electronically highlighted and located for easy identification[3]

B. Comparison of methods and approaches

In their paper titled New Features of Cervical Cells for Cervical Cancer Diagnostic System Using Neural Network. by Mustafa, N., N. A. Mat Isa et al. , it has been stated that though Pap test is the most popular and effective test for cervical cancer, Pap test does not always produce good diagnostic performance. Taking a note of this problem, they have proposed to develop diagnosis system based on neural networks to increase the diagnostic performance. The method used by them is as follows :

1) In order for neural networks to be used as cervical cancer diagnostic system, the features of cervical cell are used as inputs for neural networks.
2) classification of cervical cell type are used as output target.
3) The new cervical cell features are extracted from ThinPrep images.[4]

In the paper by Debasis Bhattacharyya et al. , Preprocessing for Automating Early Detection of Cervical Cancer. ,states that In Cervigrams , cervix region occupies about half of the raw cervigram image. Other parts of the image contain inconsequential information, such as equipment, frames, text and non-cervix tissues. This irrelevant information can muddle automatic identification of the tissues within the cervix. Therefore cervical borders are focused to have a geometric boundary on the relevant image area. This has been achieved by removal of specular reflection (SR) from raw cervigrams using region growing techniques and clustering.[5]

Asselin, Marie-Claude,et al .discuss the imaging methods available to provide appropriate biomarkers of tumor structure and function using selective regions of interest (ROI) ,Cluster analysis and Histogram analysis.[6]

In the report by Demir, Cigdem, et al., titled Automated cancer diagnosis based on histopathological images: a systematic survey. ,they have emphasized on Design of computational diagnostic tools which enable objective judgments by making use of quantitative measures as opposed to traditional cancer diagnosis, where pathologists examine biopsies to make diagnostic assessments largely based on cell morphology and tissue distribution which is subjective and often leads to considerable variability. For this, the following has been used by them :

1) determination of focal areas using image processing,
2) feature extraction to quantify the properties of these focal areas, and
3) focal areas are then classified as malignant or not or identifying their malignancy levels. [7]

In a paper titled Detection of Cancer in Pap smear Cytological Images Using Bag of Texture Features ,by Edwin Jayasingh M, Allwin.S , the problem of developing A visual dictionary based method for content based image retrieval in cervical microscopy images using texture features has been discussed. Here, The nucleus region in every image is identified by a simple and reliable segmentation algorithm and texture features are extracted from blocks of these region. These features from the entire database are clustered to build a visual dictionary.[8]

In the paper titled Cervix Cancer Diagnosis from Pap Smear Images Using Structure Based Segmentation and Shape Analysis, by Lipi B. Mahanta,et al. , an approach to automate the screening process for analysis of PAP smear images of cervical region based on cell nuclei distribution and shape and size analysis is given .The method used is histogram and structuring element based segmentation and feature extraction [9]

PROPOSED SYSTEM

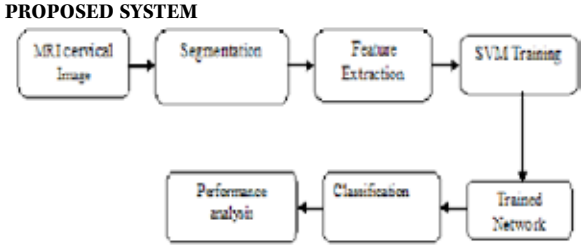

fig. 1 **proposed system**

As discussed in the section II above, the existing methods which are in practice for detection and classification may not prove fully effective in some scenarios due to the aforementioned limitations :

1) Due to the anatomy of the area , consistent and reliable sampling cannot be guaranteed.
2) cervical-cancer screening remains a laborious task in terms of the means of taking input and the detection procedure ,which is manual.
3) The procedure requires a trained colposcopist and can be expensive to perform.
4) Moreover, there is always a possibility of a screening error when cytotechnologists miss abnormal cells.
5) Accuracy associated with false negatives may be high in some cases.

An automated decision support system using magnetic resonance images for early detection is proposed in this paper that is used for diagnosis of cervical precancerous and cancerous lesions. The scheme is presented as a block diagram in Fig. 1. The process of translating magnetic resonance image into a thorough diagnosis of cancer is decomposed into four modules:

a) preprocessing of image and segmentation;
b) Feature selection and feature extraction;
c) Classification as normal or abnormal ;
d)Optimization of classification algorithms for better results.

We are, however presenting the first three modules in this paper.

1. preprocessing and segmentation

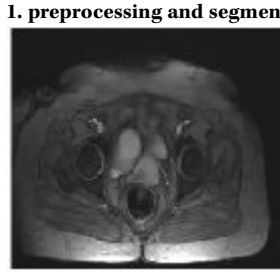

Fig 2. **Original Image**

A. Pre-processing

It is the first step in our proposed technique. the cervical MR images, which are taken as input, are subjected to be corrupted by noise during the image transmission and image digitization during image processing. The purpose of these steps is basically to covert the image to gray scale and remove the noise and improve the image quality to get more surety and ease in detecting the tumor.

B. Contrast Enhancement

Adaptive histogram equalization (AHE) is a computer image processing technique used to improve contrast in images .It differs from ordinary histogram equalization in the respect that several histograms are computed in the adaptive method, each corresponding to a specific section of the image, and uses them to reposition the lightness values of the image. It is therefore apt for bettering the local contrast of an image and bringing out more detail. Adaptive histogram equalization is an excellent contrast enhancement method for both natural and medical images. In medical images its automatic operation and effective presentation of all contrast available in the image data make it a competitor to the standard contrast enhancement methods like interactive intensity windowing.[10]

C. Morphological Operations

Dilation: The basic effect of this operator on a binary image is to progressively expand the boundaries of regions of foreground pixels (i.e. white pixels, typically).

Erosion : The basic effect of this operator on a binary image is to erode away the boundaries of regions of foreground pixels (i.e. white pixels, typically).

Filling : Closing is similar in some ways to dilation in that it tends to enlarge the boundaries of foreground (bright) regions in an image (and shrink background color holes in such regions), but it is less destructive of the original boundary shape.

D. Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments(sets of pixels or super pixels). The goal of segmentation is to ease and/or change the representation of an image into something that is more meaningful and easier to analyze.

INTENSITY BASED SEGMENTATION: One of the simplest approaches to segment an image is based on the intensity levels and is called as threshold based approach. Threshold based techniques classifies the image into two classes and works on the postulate that pixels belonging to certain range of intensity

Volume : 4 | Issue : 5 | May 2015 • ISSN No 2277 - 8179

values represents one class and the rest of the pixels in the image represents the other class. Thresholding can be implemented either globally or locally. Global Thresholding discriminates object and background pixels by comparing with threshold value chosen and use binary partition to segment the image. The pixels that pass the threshold test are considered as object pixel and are assigned the binary value 1 and other pixels are assigned binary value 0 and treated as background pixels. The threshold based segmentation techniques are inexpensive, computationally fast and can be used in real time applications with aid of specialized hardware.[11]

CLUSTERING BASED METHOD:

Fuzzy C-Means [FCM]Algorithm: In this algorithm the test pixel is allowed to be member of two or more clusters with different membership coefficient. FCM algorithm is iterative in nature and generates fuzzy partition matrix and also requires cluster centre along with objective function. The values for cluster centre and objective function are updated for every single iteration and are stopped when the difference between two successive object function values is less than some predefined threshold value.[12][13]

Fig 3. **Segmented Image using Thresholding**

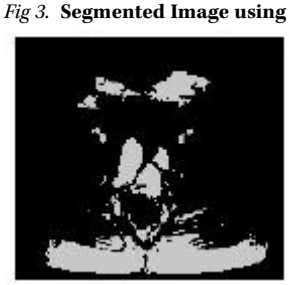

Fig 4. **Fuzzy segmented image**

2. Feature extraction

There is no explicit rationale of what constitutes a feature, and it often depends on the problem or the type of application. Feature extraction involves reducing the amount of resources needed to describe a huge set of data. The features used for the proposed system are : 'Pixel List' ,'Major Axis Length' , 'Minor Axis Length' ,'Eccentricity' ,'Orientation' and 'Area'. Also ,GLCM features like energy ,homogeneity ,correlation and entropy etc. have been used.

3. classification

Even though the paper stresses on classification using SVM, the feed forward back propagation method of ANN has also been presented in order to present comparative results and bring out the validation for opting SVM for classification.

The aim of classification is to Automatically categorize all pixels in an image into classes thereby Converting image data into information. Classification process consists of following steps: Preprocessing, Detection and extraction , Selection of the particular attribute which best describes the pattern and Classification . Classification step categorizes detected objects into predefined classes by using suitable method that compares the image patterns with the target Patterns. The following two methods have been used for classification :

ARTIFICIAL NEURAL NETWORK CLASSIFICATION

Neural networks are one of the most widely known methods. The processing elements are usually formulated into an array of layers with connections between the layers. Basically, ANN is split into three layers which are an input layer, at least one hidden layer and an output layer. they can be used to model some mapping between sets of input and output variable with appropriate pattern of weights. Back propagation learning uses the gradient descent procedure to remodel the connection weights which is derived from the consideration of minimizing some error function. This error function is needed to change the network parameter, which proves to be advantageous in improving the network performance.

Artificial Neural Network training was developed using MATLAB 2012b software. The network chosen for the pre-diction neural network had one input layer, 2 hidden layers and one output. For the hidden layers, the number of neurons is obtained by trial and error. We used cross validation method to test our classifiers. Out of 40 MR images, one image is used for testing while the remaining 39 images are used for validation. This process is continued such that within each iteration a different image of the data is held-out for validation and the rest images are used for learning.

SUPPORT VECTOR MACHINE CLASSIFICATION

SVM is one of the excellent tools for classification and regression problems with a good generalization performance. SVM formulates a hyper plain or a set of hyper plains to partition the two sets of data in a feature space. The key approach of SVM is to try finding the choicest hyper plain by maximizing the minimum margin between the two sets.

The SVM was also run by using program MATLAB 2012b. Total 40 samples are used as input signals. Linear SVM is used as kernel function for training SVM. Usually the linear kernel is much rapid in training and testing speed. An important merit of linear classification is that training and testing schemes are much more efficient. Just like neural network, the experiment was done in iterations of cross validation and the final average performance of the classifier was taken out.

RESULTS

The results for both the classifiers were found out after creating a database first. The results obtained were compared with the ground truth and for each image, the classification results were analyzed. Having obtained this information, the accuracy of the classifiers were calculated . The accuracy for SVM comes out to be 92% while that for ANN comes out to be 84%.

CONCLUSIONS

Magnetic Resonance Imaging (MRI) is an advanced medical imaging technique providing rich information about the human soft tissue anatomy. However in MRI images, the amount of data is extensive for manual interpretation and analysis and this has been one of the biggest hitches in the effective use of MRI. The goal of Magnetic Resonance (MR) image segmentation is to accurately identify the principal tissue structures in these image volumes. In the specific case of cervical MRI, the problem of segmentation and classification is particularly critical for diagnosis and treatment purposes. It is necessary to develop algorithms to obtain robust image segmentation and classification such that the end classification can be expected to reliable.

FUTURE WORK

Graphical results are proposed in order to have a better visualization of the results. Optimization of the presented algorithms

can be done in order to increase the classification efficiency. Future scope of this research is to explore whether the use of imaging techniques like MR imaging and automated image processing could provide an reasonable substitute to Pap screening and traditional colposcopy.

REFERENCE

[1] Cancer Facts and Figures 2013, available at: http://www.cancer.org. | [2] Rajeswari, S., and K. Theiva Jayaselvi. "Support Vector Machine Classification for MRI Images." International Journal of Electronics and Computer Science Engineering 1, no. 3 (2012): 1534-1539. | [3] Shanthi, S., and V. Murali Bhaskaran. "Intuitionistic fuzzy C-means and decision tree approach for breast cancer detection and classification." Eur J Sci Res 66, no. 3 (2011): 345-351. | [4] L. G. Koss, The Papanicolaou test for cervical cancer detection. A triumph and a tragedy,J. Am. Med. Assoc. vol.261, 737743 ,1989. | [5] article on thinprep from medicals of professionals. | [6] artical-New Tools Improve Diagnosis of Cervical Cancer. | [7] Mustafa, N., N. A. Mat Isa, M. Y. Mashor, and N. H. Othman. "New Features of Cervical Cells for Cervical Cancer Diagnostic System Using Neural Network." IJSSST, Vol. 9, No. 2, | May 2008, | [8] Das, Abhishek, Avijit Kar, and Debasis Bhattacharyya. "Preprocessing for Automating Early Detection of Cervical Cancer." In Information Visualisation (IV), 2011 15th International Conference on, pp. 597-600. IEEE, 2011. | [9] Asselin, Marie-Claude, James PB OConnor, Ronald Boellaard, Neil A. Thacker, and Alan Jackson. "Quantifying heterogeneity in human tumours using MRI and PET." European Journal of Cancer 48, no. 4 (2012): 447-455. [10] Demir, Cigdem, and Blent Yener. "Automated cancer diagnosis based on histopathological images: a systematic survey." Rensselaer Polytechnic Institute, Tech. Rep (2005). | [11] Edwin Jayasingh M, Allwin.S, Detection of Cancer in Pap smear Cytological Images Using Bag of Texture Features ,IOSR Journal of Computer Engineering (IOSR-JCE)Volume 11, Issue 1 (May. - Jun. 2013), PP 01-07 | [12] Lipi B. Mahanta, Dilip Ch. Nath, Chandan Kr. Nath Cervix Cancer Diagnosis from Pap Smear Images Using Structure Based Segmentation and Shape Analysis, Journal of Emerging Trends in Computing and Information Sciences, VOL. 3, NO. 2, February 2012 | [13] A. M. Khan, Ravi. S,"Image Segmentation Methods: A Comparative Study" International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-4, September 2013 | [14] "A Survey of Image Segmentation Algorithms Based On Fuzzy Clustering", IJCSMC, Vol. 2, Issue. 7, July 2013, pg.200 206 SURVEY ARTICLE | [15] Mahesh Yambal, Hitesh Gupta ,"Image Segmentation using Fuzzy C Means Clustering: A survey", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 7, July 2013 |