International Journal of Engineering Research and Technology. ISSN 0974-3154 Volume 10, Number 1 (2017)

Early Diagnosis of Endometrium Cancer using ANN- A Classifier

Mrs. Snehal R. Shinde
ME (E & TC), D. Y. Patil College of Engineering. & Technology, Kasaba Bawada, Kolhapur, Maharashtra, India

Prof. Dr.K.V.Kulhalli
H.O.D. Department of Information Technology, D. Y. Patil College of Engineering. & Technology., Kasaba Bawada, Kolhapur, Maharashtra, India

Abstract
Early diagnosis represents one of the most promising approaches to reducing the growing cancer burden. In this paper, computer aided diagnosis (CAD) system for detection of Endometrium Cancer is developed by analyzing the ultrasound images of the uterine. CAD system helps to improve the diagnostic performance of radiologists in their image interpretations. The results of every module are demonstrated in this paper. The ultrasound images that are acquired will be filtered using median filter & then segmented to obtain Region of Interest ROI using Region Growing Segmentation Algorithm. The different statistical features of the ROI will then be extracted which acts as an input to the ANN thus providing results as to Cancerous or Non-Cancerous for the tumor.

Keywords: Artificial Neural Network, Back Propagation, Benign, CAD, Malignant, Median Filter, Region Growing Algorithm.

Introduction

The most important resource of the country is its 1250 million population where there is only 1 Community Health Center (CHC) for every 2 to 3 lac population i.e. only 1% of the population. Also poor access to the health facilities in many rural areas adds up to the situation, decreasing the survival rate. Thus, in a country like India it is very essential to create awareness regarding the diseases & its symptoms.

The diagnosis of any medical image is done by humans. The accuracy of this diagnosis depends both on the quality of image acquired and its interpretation. But there are several problems with human based diagnosis; these include perception, experience, fatigue, bias, and noise. To avoid them, computer-aided diagnosis (CAD) system has become popular.

Uterine-cancer or Endometrial-cancer is one of the most common malignant gynecological tumors; the 4th most common Cancer to be specific; that prove to be a serious threat to women health. Thus, it is crucial to detect this Cancer at an early stage. Hence, by using a CAD system, a radiologist considers the output from a computer analysis of the interpreted image, though the final diagnosis is made by the radiologist.

The goal of CAD is to reduce search and interpretation errors as well as variation between and within observers. Thus, to reduce the mortality rate due to Endometrial cancer, CAD systems prove to be the best tool.

Methods

It is proposed to develop "The simulation models of Endometrial Cancer detection using MATLAB tool and different equations. The simulation results will be then evaluated with the current method's outcomes.

The projected CAD system involve following steps as shown in the Figure 1.

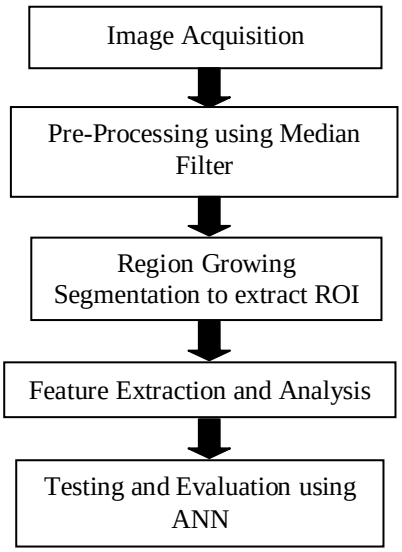

Figure 1: Block Diagram of proposed system

i) Image Acquisition: If the patient has symptoms that suggest uterine cancer, then the doctor will try to find out what's causing the problems. The patient may have to do the following test:

Ultrasound: An ultrasound device uses sound waves that can't be heard by humans. The sound waves make a pattern of echoes as they bounce off organs inside the pelvis. The echoes create a picture of your uterus and nearby tissues. The picture can show a uterine tumor. For a better view of the uterus, the device may be inserted into the vagina (transvaginal ultrasound).

Different images either malignant or benign are acquired from different medical atlases, pathological labs, internet websites & hospitals. The images that are acquired are completely unprocessed. The images can be in the form of DICOM, TIFF, JPG, PNG etc. All these images are converted into uint.8 format. Uint.8 is a format where the image is converted into

integer value so that further pre-processing can be carried out on it.

ii) Image Pre-Processing: Pre-processing technique consists of Noise removal or noise suppression. Several filters are used for noise removal that is well established in gray value image processing. Few of them are Wiener Filter, Averaging Filter & Median Filter. The results of Median Filter showed great results in comparison with the others. The figures depicted below (Fig. 2) show the results after pre-processing using Median Filter.

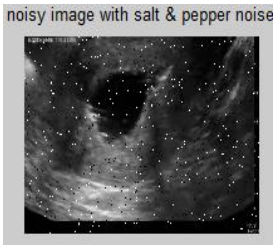

Fig.2: (a)

Fig.2: (b)

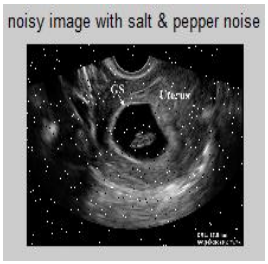

Fig.2: (c)

Fig.2: (d)

Figure 2:
2(a) Benign Image with Salt & Pepper noise;
2(b) Filtered Benign Image using Median Filter;
2(c) Malignant Image with Salt & Pepper noise;
2(d) Filtered Malignant Image using Median Filter;

Hence, Median Filter proves to be the best that improves salt & pepper noise from the image & provides an improved quality of image.

iii) Region Growing Segmentation: Image segmentation is considered to be another major step. Successful biological image analysis usually requires satisfactory segmentations to identify regions of interest as an intermediate step. Segmentation subdivides an image into its constituent regions or objects. Segmentation should stop when the objects or ROI in an application have been detected. Region Growing Segmentation approach is followed in this system which showed better results when compared with other algorithms.

The results obtained from a Benign & a Malignant Endometrium image are shown in the Figures (3) (4) & (5) below.

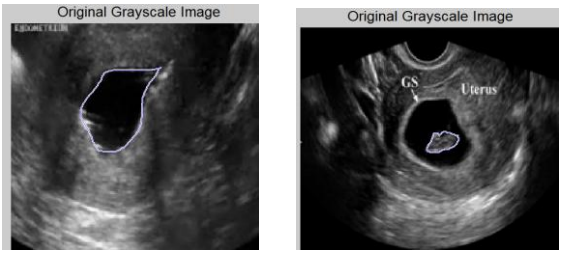

Fig.3: a) Selected ROI of Benign Image; b) Selected ROI of Malignant Image

Fig.4 : a) Area of Benign Image after 200 iterations; b) Area of Malignant Image after 10 iterations

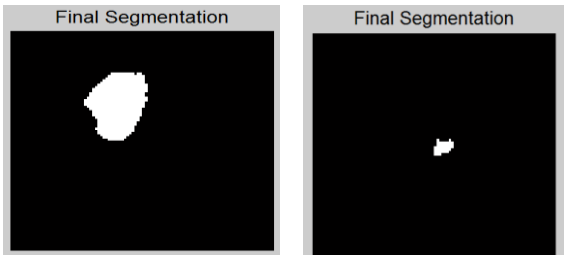

Fig.5: a) Final result of segmentation for Benign Image; b) Malignant Image

iv) Feature Extraction & Analysis of ROI: Feature extraction is a method of capturing visual content of images for indexing as well as retrieval.

After segmentation, set of features are required for each image that act as an input to the classifier. The different features such as Skewness, kurtosis, entropy, standard deviation, area, white pixel count, energy were extracted. But only a few seem to give accurate results when fed to the classifier.

Thus only 4 Statistical features were chosen & they are Area, White Pixel count, Entropy & Standard Deviation. The results of these features are illustrated in Table 1 & 2

v) Artificial Neural Network- A Classifier: Artificial neural networks are relatively crude electronic networks of "neurons" based on the neural structure of the brain. They process records one at a time, and "learn" by comparing their classification of the record with the known actual classification of the record.

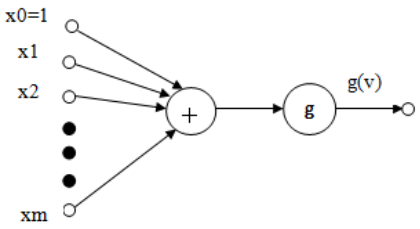

Fig.7: General ANN Architecture

The errors from the initial classification of the first record are then fed back into the network, and used to modify the networks algorithm the second time around, and so on for many iterations.

This method is known as Error back Propagation which is been implemented in the proposed system. The final result of this classifier will be either Malignant or Benign i.e. cancerous or non-cancerous respectively.

Performance Measures: Performance Parameters can be calculated using the following notations, Positive = identified & Negative = rejected. Thus,

a. **True Positive (TP) –**
   Sick people correctly diagnosed as sick.
b. **True Negative (TN) –**
   Healthy people correctly identified as healthy.
c. **False Positive (FP) –**
   Healthy people incorrectly identified as sick.
d. **False Negative (FN) –**
   Sick people incorrectly identified as healthy.

Table 1: Features Extracted from 6 Benign Samples

Benign Samples					
Samples	**Area**	**White pixel count**	**Entropy**	**Std. Deviation**	**Result**
B1	339.75	3394	7.4718	59.7769	**TN**
B2	810	4877	6.6687	53.8096	**TN**
B3	1245.875	2542	7.3022	55.757	**TN**
B4	208.125	2626	6.8683	49.5496	**TN**
B5	1962.25	18570	6.8045	71.4256	**FP**
B6	488.25	2062	6.2889	59.6927	**TN**

Table 2: Features Extracted from 10 Malignant Samples

Malignant Samples					
Samples	**Area**	**White pixel count**	**Entropy**	**Std. Deviation**	**Result**
M1	58	364	6.0838	98.6641	**TP**
M2	98.125	2607	6.6837	56.8241	**TP**
M3	223.4	2426	6.8159	72.5375	**TP**
M4	41	4936	6	63.7676	**TP**
M5	416.375	5008	7.0491	64.2064	**FN**
M6	310.375	611	4.1194	67.8407	**TP**
M7	34.5	644	6.1672	85.3009	**TP**
M8	1140.25	2080	7.2836	58.7513	**TP**
M9	146.25	1191	7.0673	73.4093	**TP**
M10	822.625	4378	7.4431	59.2278	**TP**

The summary & graphical representations of the above results are shown as below:

Table 3: Summary of Features Extracted from 16 Samples

Total no. of Test images	Total TP	Total TN	Total FP	Total FN
16	9	5	1	1
%	56.25	31.25	6.25	6.25

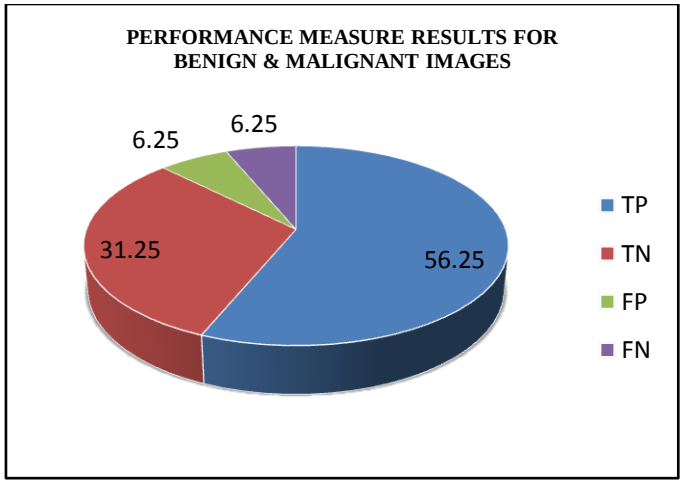

Fig.8: Graphical representation for performance measure parameters.

The important performance measure that needs to be calculated is Accuracy of the system which is given below:

i) Accuracy: Accuracy is the nearness of a calculation to the true value.

Thus, Accuracy: TP + TN/ (TP + FP + TN + FN) x 100
= 9 + 5/ (9+01+5+01) x 100
= 87.5%

ii) Sensitivity: Sensitivity is the probability that a test will indicate 'disease' among those with the disease.

Thus, **Sensitivity: = TP/ (TP + FN) x 100**
= 9/ (9+01) x 100
= 90%

iii) Specificity: Specificity is the fraction of those without disease who will have a negative test result

Thus, **Specificity: = TN/ (TN + FP) x 100**
= 5/ (5+01) x 100
= 83.33%

Discussion

From the references & personnel discussions with the researchers, professors got an important idea for detecting Endometrial Cancer using Computer vision & Image processing techniques. Endometrium cancer is the fourth most common cancer amongst women. The current methods are very lengthy & sometimes can provide false results due to various reasons of the radiologists. Hence, the CAD system helps the physician as well as the radiologist to identify the suspicious nodules of cancers & to provide about 90% correct results. The CAD Systems are beneficial to detect cancerous nodules within no time, thus increasing sensitivity, specificity, accuracy & efficiency of the diagnosis.
The most important challenge here is the diagnosis of this cancer which usually occurs at a very late stage. Hence, women should have a regular check-up with their doctors to prevent from this Cancer.

Conclusion

The proposed project of Endometrial Cancer Diagnosis using CAD systems is useful to improve the present methods of diagnosis to obtain an accuracy of 87.5%. Also the proposed project proves to be the best system in analyzing the stage at which the cancer has spread as well as its features such as area & radius. The Computer Aided Diagnostic (CAD) Systems thus are truly beneficial to detect cancerous nodules.

References

[1] Tartar, Ahmet, Akan, Aydin, "Performance of ensemble learning classifiers on malignant-benign classification of pulmonary nodules", IEEE, PP. 722-725, April 2014.
[2] Q Zheng, Yingjie liu, WeiLiang Zhu, "Uterine calcifications segmentation and extraction from ultrasound images based on level set", IEEE, Volume 2, PP. 591 – 594, Nov 2013.
[3] Thampi, L.L., Malarkhodi S., "An automatic segmentation of endometrial cancer on ultrasound images" IEEE, Volume 2, PP. 139 – 143, April 2013.
[4] Ziqian Zhang, Tiexiang Wen,Jia Gu, Lei Wang, "Cancer area extraction for uterine cervical image based on spectral matting" IEEE, Volume 3, PP. 1351 - 1354, Oct 2010.
[5] Yatharth Saraf, "Algorithms for Image Segmentation", Thesis at Birla Institute of Technology and Science, Pilani, Rajasthan, 4th May, 2006.
[6] Richard Alan Peters, "A New Algorithm for Image Noise Reduction using Mathematical Morphology" IEEE Transactions on Image Processing, Volume 4, Number 3, pp. 554-568, May 1995.