ISSN:0975-9646

Study of a Disease Sample Pixel Clustering for Accurate Infection Prediction Using Image Processing and Data Mining

Mr. Milind H. Padgavankar[#1], Mr. Shrikant P. Akarte[*2]

[#1]*ME (CSE) ,Second Year,Department of CSE,*
Prof. Ram Meghe Institute Of Technology and Research, Badnera, Amravati
Sant Gadgebaba Amravati University, Amravati, Maharashtra, India – 444701.

[*2] *Assistant Professor, Department of CSE, Prof. Ram Meghe Institute Of Technology and Research, Badnera, Amravati.*
Sant Gadgebaba Amravati University, Amravati, Maharashtra, India – 444701.

Abstract— The liver is one of the most important organs of human body. Liver diseases result heavy death worldwide. In helping doctors and surgeons treating liver diseases, computer aided liver disease diagnosis and liver surgical planning systems are playing important roles nowadays. The proposed detection methodology makes use of MRI, CT and USG scan imagery. K-means clustering technique is adopted so as to segment the images in order to capture the region of interest. Later, Haar wavelet transform is considered to compute the threshold values for the region of interest. It provides accuracy over 85%.

Key words: -Data mining, Haar wavelet transform, Image processing, k-means clustering, Liver cancer.

I. INTRODUCTION

Liver is one of the most important organs of human body. Liver diseases result heavy death worldwide. In helping doctors and surgeons treating liver diseases, computer aided liver disease diagnosis and liver surgical planning systems are playing important roles nowadays. An accurate and automatic segmentation approach of liver parenchyma, vessel and tumour is crucial to a com-puter aided liver disease diagnosis and liver surgical planning system such as a system for liver transplantation. However, due to the highly varying shape of liver, low contrast and intensity in homogeneity inside liver, weak boundaries to its adjacent organs such as heart and stomach, and intensity homogeneity to adjacent organs, liver segmentation becomes a challenging task that has attracted much research attention recently. The prime facie of this disease is due to the regular or excessive consumption of alcohol, intake of contaminated food and drugs, injecting drugs with shared needles. Apart from these, having low immunity, inherited liver diseases, L-carntine deficiency, smoking,etc.

II. LITERATURE REVIEW

Research is a process of identification of problem in a particular field by adopting any of the existing or introducing new techniques. In this paper a method is introduced to diagnose liver cancer which is considered to be novel and pragmatic. Various scientist has put forward their utmost effort in the B. Prem et al. [1] put forth a proposal for Liver surgery based on the Vascular Territories in which data related to portal and hepatic vein were used for experimentation. Smriti Sahu et al. [2] proposed an image enhancement technique for ultra sound images such as contrast stretching, shock filters and so on so as to provide an effective clinical diagnosis approach. Pradeep Kumar B.P et.al [3] proposed a fully automatic segmentation of Ultrasound liver images using Peak and valley method which is a new nonlinear, non iterative multidimensional filter for impulsive like noise reduction. S.S. Kumar, Dr. R. S. Moni et al [4] proposed a computer aided analytical system for the diagnosis of benign and malignant liver tumors from computed tomography (CT) images using curvlet transform based multiresolution texture feature extraction and neural network. Ekong V.E. et al. [5] have developed a fuzzy cluster means system to support the diagnosis of liver diseases using a set of clinical signs and symptoms with LFTs. The experimental results showed a quality enhanced liver diseases diagnosis, but with a time complexity. Robin Martin, et al. [6] implemented a semi-automatic method based on region-growing to isolate the liver part. Smitri et al, [7] uses the image enhancement technique along with filtering technique over the ultrasound images so as to enhance and read the scan image well for proper detection. Himadri Nath Moulick and Moumita Ghosh [8] presented how images can be compressed so as to reduce the memory size and how image can be clustered using k means clustering technique. Piotr Porwik and Agnieszka Lisowska [9] in their paper gave a brief on wavelet transform of which Haar wavelet transformed is elucidated. Pansnur M.A and P.S. Malge [10] represent the technique of image retrieval using Haar wavelet transform and K means clustering technique.

III. OBJECTIVE

This includes the objective of this research work is to process and analyze the images that are framed from the CT scan, MRI scan [12] and generate results whether the said image contains cancer cells or not. These images are collected from various multispecialty hospitals and diagnostic centers. The experimentation procedure makes use of MATLAB R2011a software in order to process the images. The overall methodology is elucidated by the segmentation process represented. As the images that are

framed via the MRI/CT scan exist as a grayscale image. It creates a discrepancy in identifying the cancer cell which may mislead the experimentation process, therefore the image is again converted to RGB image which makes it easier to identify the cancer cell based on the color. Now, as the cancer cell is considered to be the region of interest segmenting [18] the liver alone from the abdominal CT image is difficult due to the fact that the image includes other organs like kidney, spleen, pancreas etc very close to the liver. In order to amass only the liver part and analyze the cancer cell the experimentation makes use of image segmentation using *K*-means clustering [10]. The clusters thus formed by the above mentioned process. Now the clustered image that depicts the extracts of cancer cell which is further used for detection process. So as to analyze and judge if the given image is a cancer cell or not the feature is extracted by cropping the region the interest, and for these images the threshold range is to be fixed.

IV. METHODOLOGICAL ANALYSIS

The obtained values are hence forth substitute to find the final calculated threshold using equation 1,this threshold value is calculated using the square root of 2 multiplied with the product of variance and log(n) value which is divided to n and then multiplied to k where n represents the number of pixels which is obtained by the clustered image by area calculation and k represents the number of clusters formed.

$$T_2 = \sqrt{\frac{2*\sigma^2*\log(n)}{n}} * k \qquad (1)$$

Further the MATLAB obtained threshold value and the calculated value are compared and finally the mean of these two threshold values is calculated, using equation 2, which further results in generating a final threshold value to justify if the given clustered image carrying the affected part is indeed a cancer cells or not. Here finally images are justified by analysis and comparison method wherein, if the threshold value obtained through the software (T1), the calculated threshold value (T2) and finally the average value of the above thresholds (T3) falls within the same range with a marginal variation which can be neglected, depicts the existence of cancer cell or not.

$$T_3 = \frac{T_1 + T_2}{2} \qquad (2)$$

If all the threshold value falls within the same range the image is said to have cancer cells effected, if the variations are very much out of boundary then it is considered to be a no cancer part. The results are established and tabulated in Table I. Consider Images, its T1 value is 8.50, T2 value is 9.10 and T3 value is 8.80. All the three values fall in the same range hence the experimental result is cancer. Further, we compared the experimental result with historical data. For Image , both experimental and historical results are cancer and hence, the final prediction is true.

Image No	Number of Pixels (N)	Variance	T_1	T_2	T_3	Experimental result	Historical Data	Prediction
Image 1	15.81	5.46	8.50	9.10	8.80	cancer	cancer	TRUE
Image 2	10.28	3.72	8.81	8.56	8.69	cancer	cancer	TRUE
Image 3	14.85	3.74	7.83	7.68	7.76	cancer	cancer	TRUE
Image 4	8.36	2.88	7.38	7.97	7.67	cancer	cancer	TRUE

Table.1 value for Cluster region of Liver

V. COMPARATIVE ANALYSIS

Cancer detection is one of the challenging tasks that researchers of present day are experiencing. After a lot of research, still there is a lack of an accurate model as the detection is not unidisciplinary but is a multi-disciplinary task as it depends on various parameters. A lot of research has been performed on the development of accurate techniques for the detection of cancer at the earlier stage.

VI. RESULT

This paper this contributes by providing a computer aided diagnostic system for the diagnosis of the liver cancer using the images framed through the MRI, CT scan of certain patients. This diagnostic application makes use of MATLAB software for processing of the image, by making use of Haar wavelet transformed and clustering techniques. The whole analysis is done based on the threshold values and the images are justifying by checking if the threshold falls within the same range estimated for each image. The result obtained found to be a pragmatic approach for the early and accurate detection of cancer cells. The experimentation gave an accuracy of about 85% besides being less time complex, reducing the computational complexity for the purpose of detection.

REFERENCES

[1] P. Sheba lin, V. Keilis - Borok , A. Gabri elov, I. Zaliapin, D.Turcotte, “Short-term earthquake prediction by reverse analysis of lithosphere dynamics,” *ELSEVIER Tectonophysics* , 2006, pp.63 – 75

[2] Neeti Bhargava, V. K. Katiyar, M. L. Sharma and P. Pradhan, “Earthquake Prediction through Animal Behaviour: A Review,” *Indian Journal of Biomechanics*: Special Issue NCBM 2009, pp.159-165.

[3] G. Molchan and L. Romashkova, “Earthquake prediction analysis: The M8 algorithm,” *physics.geo-physics*, 2010, pp. 1-20

[4] Sajjad Mohsin, and Faisal Azam, “Computational seismic algorithmic comparison for earthquake prediction,” *International Journal Of Geology* Issue 3, Volume 5, 2011, pp. 53-59.

[5] Chieh-Hung Chen, Horng-Yuan Yen,Chung-Ho Wang, Yih-Hsiung Yeh, Jann-Yenq Liu, Yee-Ping Chia, Chen Liu, Yetmen Wang and Wen-Tzong Liang, “Identification of earthquake signals from groundwater level records using the HHT method,” *Geophysical Journal International*, 2010, pp.1231–1241.

[6] D. Giaouris, J.W. Finch, “Denoising using wavelets on electric drive applications,” *ELSEVEIR*, Electric Power Systems Research, 2008, pp.559–565.

[7] Radomir S. Stankovic and Bogdan J. Falkowski “The Haar wavelet transform: its status and achievements,” *ELSEVEIR*, Computers and Electrical Engineering, 2003, pp.25–44.

[8] Burhan Ergen, “Signal and Image Denoising Using Wavelet Transform,” Advances in Wavelet Theory and Their Applications in Engineering, Physics and Technology, 2012, pp.495-514.

[9] Piotr Porwik and Agnieszka Lisowska, “The Haar–Wavelet Transform in Digital Image Processing:Its Status and Achievements,” *Machine Graphics and Vision*, 2004, pp.79-98.

[10] Dr. Micheal Sek, “Frequency Analysis Fast Fourier Transform, Frequency Spectrum,” Victoria University, pp.1-12.