International Journal of Engineering Research in Electronic and Communication Engineering (IJERECE) Vol 2, Issue 12, December 2015

Image Mining Process for Biopsy Images

[1]Ms. Nikita Ramrakhiani,[2]Mrs. Shoba Krishnan
[1]Student,[2]Professor.
[1]Master of Engineering, Department of Electronics & Telecommunication,
Vivekanand Education Society's Institute of Technology, Chembur, Mumbai, India
[2]Department of Electronics & Telecommunication, Vivekanand Education Society's
Institute of Technology, Chembur, Mumbai.

Abstract- — Biomedical imaging is becoming increasingly important as an approach to synthesize, extract and translate useful information from large multidimensional databases accumulated in research frontiers such as functional genomics, proteomics, and functional imaging. Image mining is rapidly gaining attention in the field of data mining, information retrieval and multimedia databases because of its potential in discovering useful image patterns based on color, texture, shape and basic descriptors of any image. Image mining system can help in reducing the time lag with the results as well as dependency on observations by naked human eye. This paper explains the algorithm for detecting pathogenic cells in renal tissue.

IndexTerms- Biomedical Imaging, Image Mining, Image Mining Application, Neural networks

I. INTRODUCTION

Image mining is the process of extracting previously unknown knowledge and detecting intersecting patterns from a massive set of images. It deals with the extraction of implicit knowledge, image data relationship, or other patterns not explicitly stored in the images. Thus, image mining is a collaboration of patterns. It is an interdisciplinary venture that essentially draws upon expertise in artificial intelligence, computer vision, content based image retrieval, database, data mining, digital image processing and machine learning [1]. Image mining is rapidly gaining attention in the field of data mining, information retrieval and multimedia databases because of its potential in discovering useful image patterns based on color, texture, shape and basic descriptors of any image. This technique when applied on biopsy images can give minute details and features which can detect malignancies easily.

A biopsy image gives every possible detail about an organ from the cellular level to the whole-organ level. Biomedical imaging is becoming increasingly important as an approach to synthesize, extract and translate useful information from large multidimensional databases accumulated in research frontiers such as functional genomics, proteomics, and functional imaging. [2] To fulfill this approach Image Mining can be used.

.

Image mining is rapidly gaining attention in the field of data mining, information retrieval and multimedia databases because of its potential in discovering useful image patterns based on color, texture, shape and basic descriptors of any image [10]. Image Mining will bridge this gap to extract and translate semantically meaningful information from biomedical images and apply it for testing detecting any anomaly in the target organ. [12] The essential component in image mining is identifying similar objects in different images and finding correlations in them. Integration of Image Mining and Biomedical field can result in many real world applications.

If a system is able to automatically extract relevant features directly from the images stored in the database, image database retrieval can be done efficiently. So image mining proves to be efficient as it deals with complex operations like image retrieval, indexing and storing. With image mining we will consider the four broad areas [1, 10]:

1. Finding associations
2. Classification
3. Sequential patterns
4. Time series patterns

II. MERGER OF IMAGE MINING AND BIOPSY IMAGES

There is an increase in incidence of health issues which are very difficult to diagnose, especially in developed countries. These diseases can prove fatal if not treated in time. There are few diseases whose etiologies are not clear and neither are the reasons for the increased number of cases. Delay in detecting these diseases increase complications in treating them. Existing techniques of Biomedical Imaging do not provide with immediate results. The current practice in

biomedical imaging involves presenting a 2D/3D image data after suitable processing to a human who carries out a qualitative assessment based on expert judgment.[13] Thus, these results are solely dependent on human interpretations of the biomedical images. Image mining system can help in reducing the time lag with the results as well as dependency on observations by naked human eye. Integration of bio-medical imaging and image mining makes diagnostic process efficient and swift. [2]. Detecting malignant tissue or cells at earliest stage can go a long way for successful treatment. Thus, collaboration between these two fields can nullify various procedural and clinical variations that affect the diagnose.

III. PROCESS FOR IMAGE MINING

Image mining of biopsy images deals with the extraction of implicit knowledge, image data relationship, or other patterns not explicitly stored in the biopsy images and between input image of the patient and other alphanumeric data. A minute change in a pattern in the input image is of significance in biopsy images. Low level computer vision and image processing techniques cannot be relied upon for images of such importance. [9] The basic flow of the process can be seen in Fig.1.

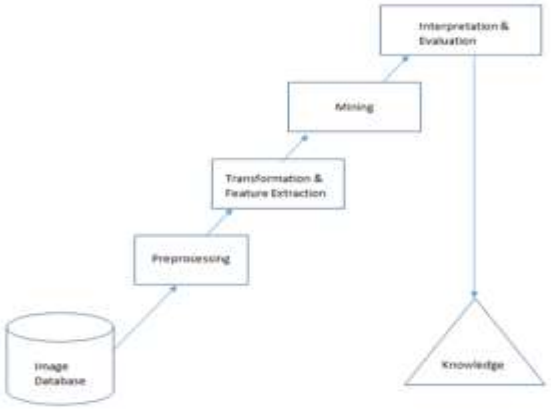

Fig.1 Process for Image Mining

In this paper, image mining algorithm is applied on renal tissue cells to detect any malignancies.

IV. PREPROCESSING TECHNIQUES FOR IMAGE MINING ON BIOPSY IMAGES

Preprocessing involves improvement of image data that suppresses unwanted distortions or enhances image features which are significant for further processing and analysis. Image pre-processing can significantly increase the reliability of an optical inspection. Preprocessing operations intensify or reduce certain image details enable an easier and better evaluation [6,8]

Initially, renal biopsy samples were stained with Sirius Red technique and image is captured with an appropriate lens. These images are then converted to grayscale images which is further divided into three planes. Preprocessing Preprocessing techniques performed were Histogram equalization Segmentation, Thresholding and Interpolation. Gray scale modification did not enhance biopsy images satisfactorily whereas thresholding, segmentation and interpolation enhanced required features in the same. Result for histogram equalization can be seen below:

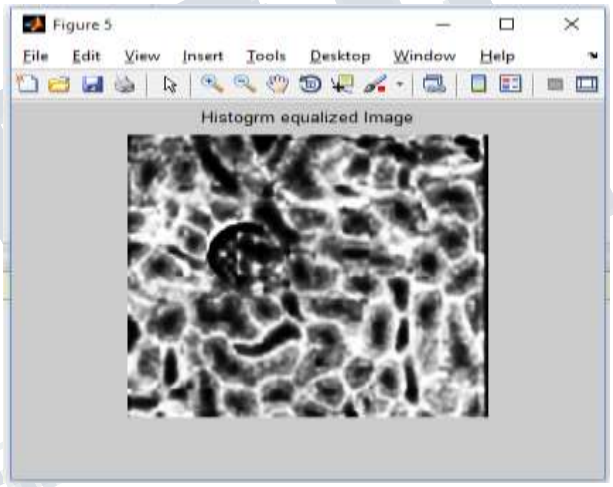

Fig.2 Histogram Equalized Image

V. SEGMENTATION SQUARE CHARACTERIZATION

After preprocessing, segmentation was performed using Segmentation Square Characterization method where we divide the image into 37x37 size blocks. This value has been decided to achieve a balance between accuracy and processing time. These blocks are known as Segmentation Squares. This process enables to identify the edges of the targeted area which are not clear in the biopsy images. These squares will be clustered and classified further to detect Regions of Interest (ROIs) in the biopsy images. [11] These regions vary depending on the tissue biopsied, i.e. Regions of Interest of lungs will vary from that of kidney or brain. However, basic criteria, related to the cell nuclei, will remain the same. Location of the desired cell, or shape/ structure of the cell contribute in deciding these regions of Interest.

After dividing the image in 37 x 37 pixel blocks, features are extracted from each of these and then these parts are combined to form a block for a complete detailed extraction

of edges as shown in fig. 6.2 (e). After extracting these features from all the test images, they are given as input for training neural network. Back propagation algorithm uses assigned weights for voting the majority classifiers from the training set. These weights are calculated using values of mean, contrast, energy, entropy, correlation, standard deviation and inverse difference moment. Parameters like skewness and centroid were calculated but did not give accurate results for biopsy images. Using the majority voted features and their value; preprocessed image is then segmented accordingly. Results for segmentation, block feature extraction and majority voting can be seen in fig 1.2 (a), (b) and (c).

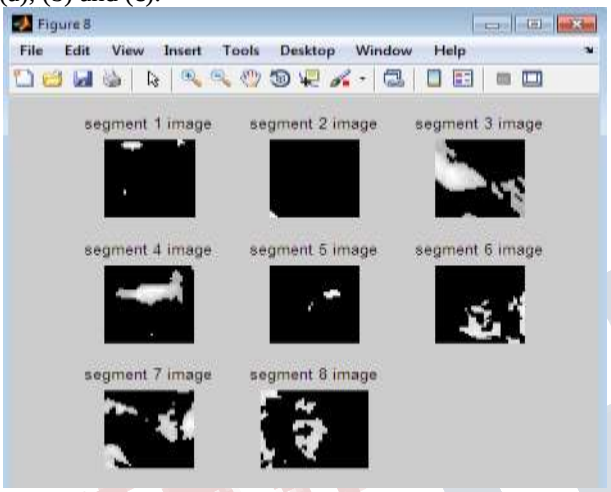

Fig.3 Segmented images of the renal biopsy cell

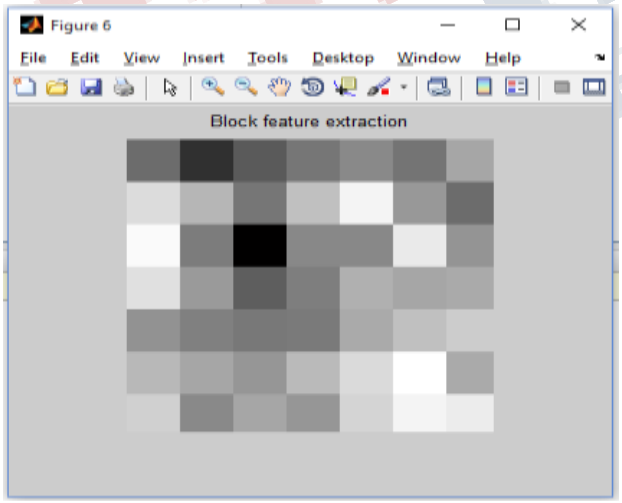

Fig4. Block Feature Extraction

As soon as aforementioned process was completed, an annotation was made on the original input image based on borders of the segmented image formed after majority voting of classifiers. This image is represented in fig. 1.2

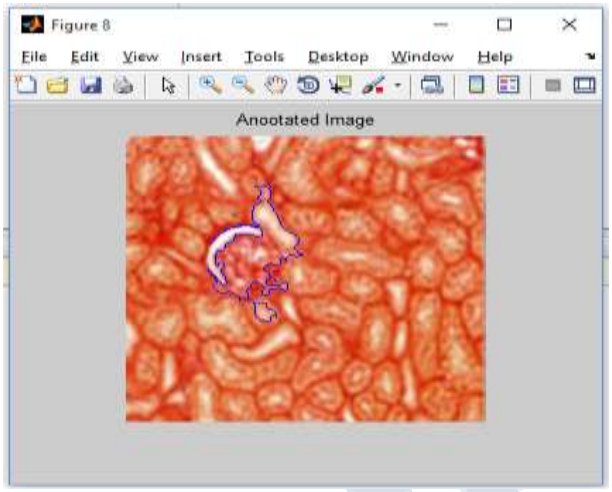

Fig5. Annotated Image

Once original image is annotated based on majority voting using neural network, based on the extracted features and the training set, result is given as pathogenic or non pathogenic. Accuracy was tested for four control images as well as test images.

VI. RESULTS

In the initial stage of the project, preprocessing techniques like Bicubic Interpolation, Cubic Interpolation on renal tissues were applied and results were analysed. With this analysis, further in the second stage of the project a database is created for these images which acted as training set for the neural network. It was concluded from the initial stage that efficiency of preprocessing techniques vary from biopsy samples. For the second stage, a total of 50 samples were taken for an accurate training set. Features like mean, contrast, energy, entropy, correlation, standard deviation and inverse difference moment were calculated and an average value was considered for higher efficiency. It can be concluded that if the average value of these 8 features is greater than 3, sample is pathogenic. Also, to remove any possible variations in the result threshold values are set for individual features as well. Values of the features calculated for four control and pathogenic samples are given below:

Image Identification	Energy	Entropy	Mean	Standard Deviation	Average value
Control Image 1	62	70	2.091	4.06	3.19
Control Image 2	73	55	1.32	3.388	3.11
Control Image 3	57	81	2.63	4.4	3.3
Control Image 4	45	14	2.76	4.5	2.81
Pathogenic Image 1	82	40	8.38	2.77	3.03
Pathogenic Image 2	90	25	4.9	2.16	3.04
Pathogenic Image 3	84	38	7	2.55	3.01
Pathogenic Image 4	87	62	7.41	3.48	3.104

Table 1. Features for control and pathogenic images

CONCLUSION

For higher accuracy in the results, features like minimum Grayscale value, maximum Grayscale value, mode, ROI's height, ROI's width, ROI's percentage, area (in pixels), median, kurtosis, skewness, Histogram's minimum value, Histogram's maximum value, area fraction, centroid, angle, and center of mass can be developed. Also, this algorithm can be modified for organ specific cells with greater database to reduce the time lag as well as procedural variations.

REFERENCES

[1] H. Wynne, L.L Mong, and J. Zhang, "Image mining: trends and developments. Journal of Intelligent Information Systems", 2002

[2] http ://www.nps.org.au/medical-tests/medical-imaging/for-individuals/advantages-and-disadvantages

[3] T. Y. Gajjar, N. C. Chauhan, "A Review on Image Mining Frameworks and Techniques", International Journal of Computer Science and Information Technologies

[4] Prabhjeet Kaur, Kamaljit Kaur, "Review of Different Existing Image Mining Techniques", International Journal of Advanced Research in Computer Science and Software Engineering Research Paper, Volume 4, Issue 6, June 2014

[5] A.Kannan, Dr.V.Mohan, Dr.N.Anbazhagan "An Effective Method of Image Retrieval using Image Mining Techniques", The International journal of Multimedia & Its Applications (IJMA) Vol.2, No.4, November 2010

[6] Er.Jyori Rani, Er.Sarabjeet Kaur, "Image Restoration Using Various Methods and Performance Using Various Parameters", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 1, January 2014

[7] Anthony Parker, Robert V Kenyon, Donald E. Troxel. "Comparison of Interpolating Methods for Image Resampling", IEEE Translations on Medical Imaging

[8] Mehmet Sezgin, Bulent Sankur, "Survey over image thresholding techniquesand quantitative performance evaluation", Journal of Electronic Imaging, 13(1), 146–165 (January2004).

[9] Ordonez, E. Omiecinski,"Image mining: A new approach for data mining," Technical Report GITCC-98-12,Georgia Institute of Technology, College of Computing, 1998

[10] D.N.D.Harini, Dr.D.Lalitha Bhaskari, "Image Mining Issues and Methods Related to Image Retrieval System", International Journal of Advanced Research in Computer Science, Volume 2, No. 4, July-August 201

[11] Theodosis Goudas, Charalampos Doukas,*A* Collaborative Biomedical Image Mining Framework: Application on the Image Analysis of Microscopic Kidney biopsies, IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 17, NO. 1, JANUARY 2013

BOOK REFERENCES

[12] Data Mining in Biomedical Imaging, Signaling, and Systems**,** Sumeet Dua, Rajendra Acharya U

[13] Medical Imaging Physics, William R. Hendee, E. Russell Ritenour.

[14] The Computational Challenges of Medicals Imaging, Michael Brenner, Alvin Despain, Robert Hederson, Darrel Long, William Press, John Tony, Peter Weinberge.