A proposed image processing framework to support Early liver Cancer Diagnosis

Aymn E.Khedr[1] and Abd El-Ghany A. M. Mohmed[2]

[1]Information system Department, Faculty of Computers and Information Helwan University, Egypt
[2]Faculty of Science, Helwan university, Egypt
Ayman_Khed@Helwan.edu.eg, Aghany72@yahoo.com

Abstract: Image recognition mining deals with the extraction of image patterns from a large collection of images stored in particular multimedia databases. Image mining is different from low-level computer vision and image processing techniques because the focus of image mining is in extraction of patterns from large collection of images, whereas the focus of computer vision and image processing techniques is in understanding and/or extracting specific features from a single image. Although there looks like to be some overlaps between image mining and content-based retrieval (both are dealing with large collection of images), image mining goes beyond the problem of retrieving relevant images. In image mining, the aim is the discovery of image patterns that are considerable in a given collection of images. Medical images include huge amount of unseen information that exploited by physicians in making reasoned decisions about a patient. However, extracting this relevant hidden information is a critical first step to their use. For this reason we use data mining techniques for efficient knowledge extraction. The dual reading (reading by two physicians or radiologists) of liver x-ray improved the accuracy rate, but at high costs. According to the fact that the medical domain involves high accuracy and particularly the rate of false negatives are very low. The computer diagnosis systems are necessary to support the medical staff to achieve high capability and effectiveness. This is the main reason for the development of classification systems to Diagnosing liver Cancer. The aims of this paper is pointed out as use some image mining techniques such as neural networks and association rule mining techniques to detection early liver Cancer using and helping physicians to decide an important decision on a particular patient state.
[Aymn E.Khedr and Abd El-Ghany A. M. Mohmed. **A proposed image processing framework to support Early liver Cancer Diagnosis.** *Life Sci J* 2012;9(4):3808-3813]. (ISSN: 1097-8135). http://www.lifesciencesite.com. 566

Key words: Classification, medical imaging, association rule mining, neural networks, image mining, Image Recognition, Early Cancer Diagnosing.

1. Introduction

With advanced in secondary storage capacity coupled with a relatively low storage cost more and more non-slandered data is being accumulated. One category of non-slandered data is image, video, sound, etc.... There are big collections of image data that can be mined to discover new and valuable knowledge. The central research issue in image mining is to how to reprocess image set or how to represents image sets in the form that supports the application of data mining algorithms. A common representation is that of feature vectors or each image representation in vector. We segment the image or this vector into sub images to finding the region of interest in image and store this region attributes in vector.

The image mining technique deals with the extraction of implicit knowledge and image with data relationship or other patterns not explicitly stored in the images. It is an extension of data mining to image domain. The main objective of this research is to apply image mining in the domain such as Magnetic Resonance Image (MRI) and/or other rays kinds to classify and detect the cancerous tissue.

Image mining of medical images is used to collect effective models, relations, rules, abnormalities and patterns from large volume of images. This procedure can accelerate the diagnosis process and decision-making. Different methods of image mining have been used to detect and classify anomalies in mammogram images such as wavelets [9], statistical methods and most of them used feature extracted using image processing techniques. Some other methods are based on fuzzy theory [6] and neural networks [3]. Here the overview of historical studies in this field:

2-background study

Ashraf et.al.(2008) [1] roposed a segmentation algorithm operates by first extracting regions satisfying the statistical characteristics (gray level distributions) of the Corpus Callosum that have relatively high intensity values. This is then processed using graph analysis and classification procedures.

Wynne et.al.(2000) [10] roposed an system called IRIS, an Integrated Retinal Information system, has been built up to afford medical professionals trouble-free and unified access to the

screening, trend and development of diabetic-related eye diseases in a diabetic patient database.

Rajendran et.al. (2010) [8] Proposed a method deals with the detection of brain tumor in the CT scan brain images. First preprocessing technique applied on the images eliminates the inconsistent data from the CT scan brain images. Then feature extraction process is applied to extracts the features from the brain images. A Novel Fuzzy Association Rule Mining (NFARM) applied on the image transaction database contains the features extracted from the CT scan brain images. The NFARM gives the diagnosis keywords to physicians for making a better diagnosis system.

Jaba et.al.(2007) [5] proposed a system using image mining techniques to categorize the images either as normal or abnormal and then classify the tissues of the abnormal brain MRI to identify brain related diseases.

Rajendran et.al.(2010) [7] proposed a method concerned with the classification of brain tumor in the CT scan brain images. The main steps involved in the system are: pre-processing, feature extraction, association rule mining and hybrid classifier. The pre-processing is done using the median filtering and edge features are extracted using canny edge detection technique. The combination of two image mining approach is been proposed. The frequent patterns from the CT scan images are produced by frequent pattern tree (FP-Tree) algorithm that mines the association rules. The decision tree method is used to categorize the medical images for diagnosis. This system improves the classification process to be more accurate. The hybrid method enhances the efficiency of the proposed method than the traditional image mining methods.

Aswini et.al. (2010) [2] applied image mining in the domain such as breast mammograms to classify and detect the cancerous tissue. A hybrid approach of feature selection using fast branch and bound algorithm and a hybrid genetic algorithms are used which approximately reduces 75% of the features and new decision tree is used for classification and provide promising results.

The objectives of the paper

The aims of this paper is pointed out as use some image mining techniques such as neural networks and association rule mining techniques to detection early liver Cancer using and helping physicians to decide an important decision on a particular patient state.

Research problem

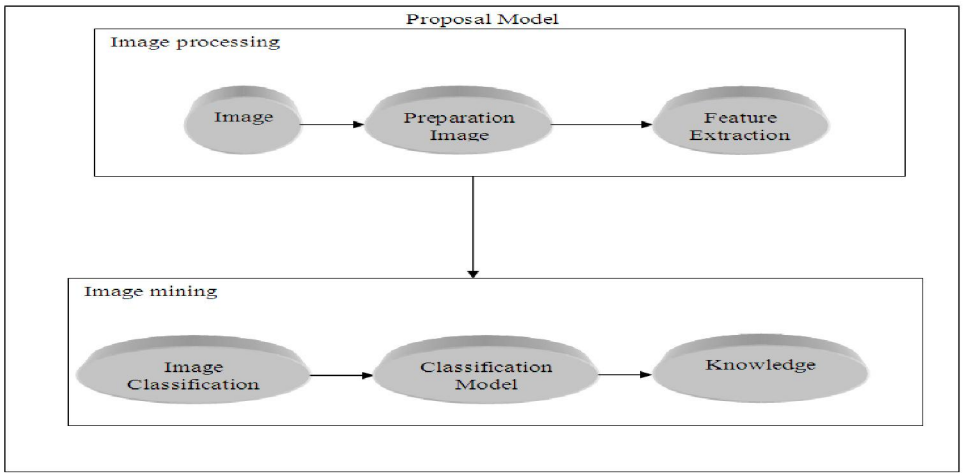
The computer aided diagnosis systems are necessary to support the medical staff to achieve high capability and effectiveness. This is the main reason for the development of classification systems to Diagnosing liver Cancer.

Image recognition mining deals with the extraction of image patterns from a large collection of images stored in particular multimedia databases. Image mining is different from low-level computer vision and image processing techniques because the focus of image mining is in extraction of patterns from large collection of images, whereas the focus of computer vision and image processing techniques is in understanding and/or extracting specific features from a single image. Although there looks like to be some overlaps between image mining and content-based retrieval (both are dealing with large collection of images), image mining goes beyond the problem of retrieving relevant images. In image mining, the aim is the discovery of image patterns that are considerable in a given collection of images [4].

Medical data mining is a promising area of computational intelligence applied to an automatically analyze patients' records aiming at the discovery of new knowledge potentially useful for medical decision-making. Induced knowledge is anticipated not only to increase accurate diagnosis and successful disease treatment, but also to enhance safety by reducing medication-related errors.

Then we can formulate the research problem as statements as following:

"This research analyze automatically the patient records to discover new knowledge"
"It increases the accuracy of diagnoses and decision making for the physicians and radiologist"



There are two tasks in our framework:

I-Image processing task

Preprocessing phase of the images is necessary to improve the quality of the images and make the feature extraction phase more reliable. This phase consists of some processes. These processes contain data normalization, data preparation, data transformation, data cleaning, and data formatting.

Using wavelet transforms represent to transform the image. We also use Segmentation to recognize regions of interest (ROI).because the medical image is large and contain huge amount of information then we use the segmentation step finds consequent regions within an image.

Images usually have a huge number of features. It is important to recognize and extract interesting features for an exacting task in order to decrease the complexity of processing. Image processing algorithms used, which automatically extract image attributes such as local color, global color, texture, and structure.

Texture is the mainly useful description property of an image and it specifies attributes, such as resolution, which used in image mining Jaba et.al.(2009) [12].

Feature extraction from images are required for many image mining applications such as content based information retrieval (CBIR), image classification etc. These features typically extracted based on the image's information by image processing only Tianxia et.al. (2008) [13].

This step consists of several steps - improve the quality of image to extracts the feature for image easily following steps:

1- Image normalization (putting the image in fixed size)

2- Image preparation (put the image in known format such as .jpg or bmp)

3- Image transformation (for example using wavelet transform)

4- Cleaning the image from noise

5- Segmentation step in which we segment the image into regions to extract the feature

6- Feature extraction

We use several feature such as:

a- Color feature is the combination of (red, blue, green) components

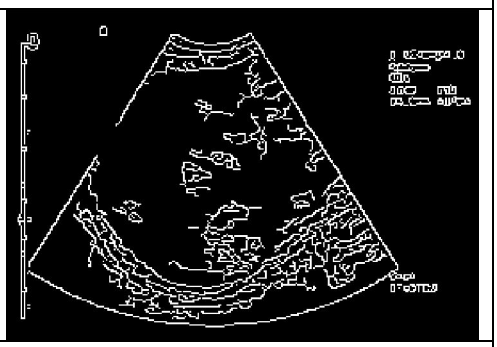
b- Texture feature: Texture is a very interesting image feature that has been used for characterization of images, with application in content based image retrieval. A major characteristic of texture is the repetition of a pattern or patterns over a region in an image. The elements of patterns are sometimes called textons.

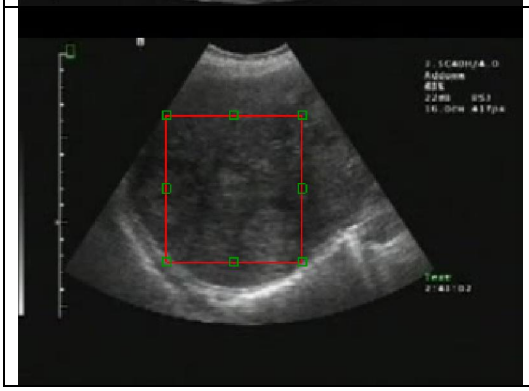
C- Edge feature is simply a large change in frequency

There are several states in liver diseases such as: Cirrhosis, Hemangioma, Hydatid liver disease, Metastases, Portal hypertension, Accessory fissure of the liver, acute calculus cholecystitis with pericholecystic abscess.

Here are some image samples and its suggested feature as mean, Standard Deviation and Variance.

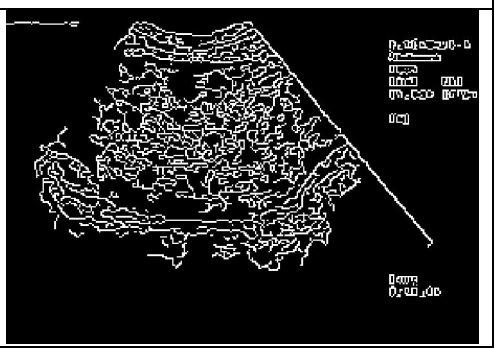
Metastases

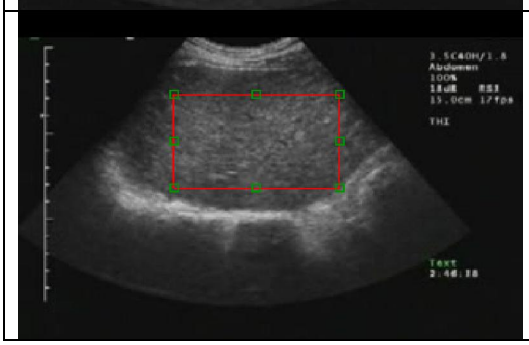




Mean: 46.4783
Standard Deviation: 12.7864
Variance: 163.493

Cirrhosis





Mean: 80.0554
Standard Deviation: 12.6949
Variance: 161.161

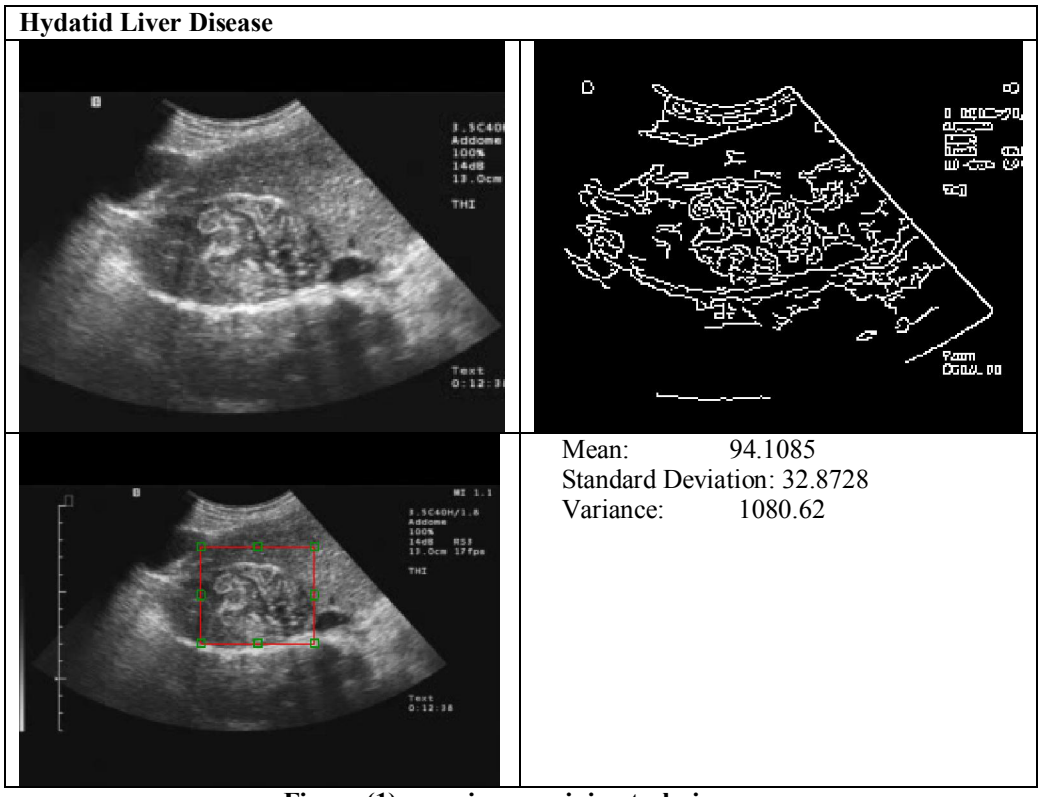


Figure (1) some image mining techniques.

II Image Mining Task

We use some image mining techniques such as neural networks and association rule mining techniques to detection early liver Cancer using and helping physicians to decide an important decision on a particular patient state. In recent years, many advanced classification approaches, such as neural networks, fuzzy-sets, and expert systems, have been widely applied for image classification.

Neural networks

Neural networks are of particular interest because they offer a means of efficiently modeling large and complex problems in which there may be hundreds of predictor variables that have many interactions.(Actual biological neural networks are incomparably more complex.) Neural nets may be used in classification problems (where the output is a categorical variable) or for regressions (where the output variable is continuous)(By Two Crows Corporation)(2008) [11].

A neural network (Figure 2) starts with an input layer, where each node corresponds to a predictor variable. These input nodes are connected to a number of nodes in a hidden layer. Each input node is connected to every node in the hidden layer. The nodes in the hidden layer may be connected to nodes in another hidden layer, or to an output layer. The output layer consists of one or more response variables (By Two Crows Corporation)(2008) [11].

Findings discussion and conclusion

This is the pilot study about using image mining for early digenesis on liver cancer. First we use image processing techniques for pre-processing and extract feature from it. Second we use image mining technique for discover knowledge from the extracted image. We used about 20 images for liver and the results is promised

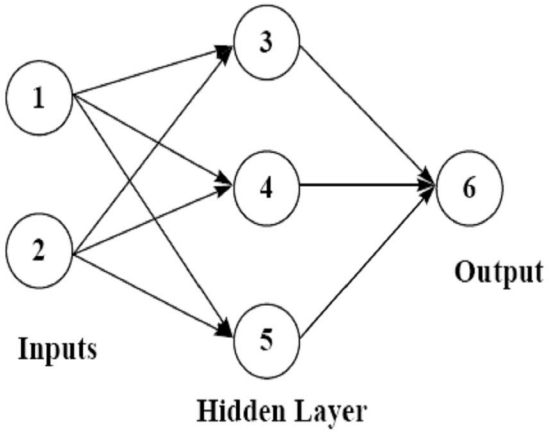


Figure (2) a neural network with one hidden layer.

References

[1] Ashraf Elsayed, Frans Coenen1, Marta García-Fiñana and Vanessa Sluming "Segmentation for Medical Image Mining: A Technical Report" 2008

[2] Aswini Kumar Mohanty, Saroj Kumar Lenka, "Efficient Image Mining Technique for Classification of Mammograms to Detect Breast Cancer" Special Issue of IJCCT Vol. 2 Issue 2, 3, 4,2010

[3] A. Dhawan et al. "Radial-basis-function-based classification of mammographic microcalcifications using texture features". In Proc. of the 17th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol 1, pages 535– 536, 1995.

[4] Ji Zhang Wynne Hsu Mong Li Lee Image Mining: Issues, Frameworks and Techniques, Page1.

[5] L. Jaba Sheela, Dr. V.Shanthi, "Image Mining Techniques for Classification and Segmentation of Brain MRI Data", JATIT, page 115-121– 2007.

[6] D. Brazokovic and M. Neskovic. Mammogram screening using multi resolution-based image segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, vol 7, issue 6, 1993.

[7] Rajendran P., M.Madheswaran "Hybrid Medical Image Classification Using Association Rule Mining with Decision Tree Algorithm" Journal Of Computing, Vol 2, Issue 1 2010

[8] Rajendran P., M. Madheswaran " Novel Fuzzy Association Rule Image Mining Algorithm for Medical Decision Support System" International Journal of Computer Applications (0975 - 8887) Vol. 1 Issue 20, 2010

[9] Wang T. and N.Karayaiannis, "Detection of micro calcification digital mammograms using wavelets", IEEE Trans. Medical Imaging, 17(4):498-509,1998.

[10] Wynne Hsu, Mong Li Lee, Kheng Guan Goh "Image Mining in IRIS: Integrated Retinal Information System"

[11] By Two Crows Corporation "Introduction to Data Mining and Knowledge Discovery" .Third Edition,2005.

[12] Jaba Sheela L and Dr.V.Shanthi "An Approach for Discretization and Feature Selection Of Continuous-Valued Attributes in Medical Images for Classification" Learning. International Journal of Computer Theory and Engineering, Vol. 1, Issue 2, page: 154, 2009 .

[13] Tianxia Gong, Chew Lim Tan, Tze Yun Leong, Cheng Kiang Lee, Boon Chuan Pang, C. C. Tchoyoson Lim, Qi Tian, Suisheng Tang, Zhuo Zhang, "Text Mining in Radiology Reports", ICDM '08. Eighth IEEE International Conference on, Page 815-820, 2008.

11/21/2012