Efficient Sonographic Image Mining Technique To Identify Acute Appendicitis

Balu Ramu *, Devi Thirupathi**
*(Ph.D Research Scholar, Department of Computer Applications,
Bharathiar University, Coimbatore, Tamilnadu, India
** (Associate Professor & Head i/c, Department of Computer Applications,
Bharathiar University, Coimbatore, Tamilnadu, India

Abstract

The acute appendicitis necessitates emergency abdominal surgery and in the past decade, sonography has gained acceptance for examining patients with acute abdominal pain. Sonographic imaging is dynamic, noninvasive, rapid, inexpensive and readily accessible. Manual analysis of sonographic image is a tedious process and consumes enormous time. The authors aim at design and development of an automatic system that would detect acute appendicitis by taking sonographic images as input. This paper describes the image mining system that automates the diagnosis of acute appendicitis with significant time reduction. The experimentation methods, results of the testing using real data are detailed in this paper. The data set of 21 patients' sonographic images collected from a reputed hospital in India has been used as input. It is concluded that an algorithm integrating region based segmentation and euclidean distance method yields accurate results in diagnosing appendicitis.

Keywords – Data Mining, Euclidean Distance, Manhattan Distance, Ultrasound, Appendicitis

I. INTRODUCTION

Appendicitis is sudden onset of inflammation of the appendix, which is a small, finger-shaped blind-ending sac that branches off the first part of the large intestine. Except for a hernia, acute appendicitis is the most common cause in the USA of an attack of severe, acute abdominal pain that requires abdominal operation. The incidence of acute appendicitis is around 7% of the population in the United States and in European countries. In Asian and African countries, the incidence is probably lower because of the dietary habits of the inhabitants of these geographic areas. Appendicitis occurs more frequently in men than in women, with a male-to-female ratio of 1.7:1. Appendicitis can affect any age but is more common in young people between 8 and 14 years. Rare cases of neonatal and prenatal appendicitis have been reported [1].

A. Related Work

Prabhudesai, Gould and Rekhraj proposed that artificial neural networks can be an useful aid in diagnosing acute appendicitis [1]. Sivasankar, Rajesh and Venkateswaran proposed that Back propagation Neural Network and Bayesian Based Classifier can be useful aid in Diagnosing Appendicitis [2]. Balu and Devi proposed that Identification of Acute Appendicitis Using Euclidean Distance on Sonographic Image [3]. Mesut Tez and Selda Tez proved that neuro fuzzy systems can incorporate data from many clinical and laboratory variables to provide better diagnostic accuracy in acute appendicitis [4].

B. This work and Major Contribution

Image mining deals with the extraction of image patterns from a large collection of images [18]. Image mining is more than just an extension of data mining to image domain. It is an interdisciplinary endeavor that draws upon expertise in computer vision, image processing, image retrieval, data mining, machine learning, database, and artificial intelligence [5]. The main advantage of using distance measures in decision making is that authors can compare the alternatives of the problem with some ideal result. Then, by doing this comparison the alternative with a closest result to the ideal is the optimal choice. Euclidean distances are special because they conform to our physical concept of distance [6]. Appendectomy remains the only curative treatment of appendicitis. The surgeon's goals are to evaluate a relatively small population of patients referred for suspected appendicitis and to minimize the negative appendectomy rate without increasing the incidence of perforation. Author's main goal is approaching 100% sensitivity for the diagnosis in a time, cost and efficient method [7].

C. Paper Organization

This paper makes two contributions; first authors find the distance between two points. Second, authors introduce a euclidean and manhattan distances. The rest of this paper is organized as follows: Section II presents overview of image mining; Section III describes system description; Section IV introduce proposed algorithm; Section V represents euclidean distance; Section VI demonstrates manhattan distance; Section VII shows experiments and results; Section VIII represents performance evaluation. Finally, Section IX concludes the paper.

II. IMAGE MINING

Image mining is a technique which handles the mining of information, image data association, or additional patterns not unambiguously stored in the images [8]. It utilizes methods from computer vision, image processing, image retrieval, data mining, machine learning, database, and artificial intelligence. Rule mining has been implemented to huge image databases [9].

The main intention of image mining is to produce all considerable patterns without any information of the image content, the patterns types are different. They could be classification patterns, description patterns, correlation patterns, temporal patterns and spatial patterns. Image mining handles with all features of huge image databases which comprises of indexing methods, image storages, and image retrieval, all regarding in an image mining system [10-11].

Rajendran and Madheswaran discussed an improved image mining technique. An enhanced image mining technique for brain tumor classification using pruned association rule with MARI algorithm is presented in their paper [12]. Dubey illustrated about an Image mining methods which is dependent on the color histogram, texture of that image. The query image is considered, then the color histogram and texture is created and in accordance with this the resultant image is found [13]. Ramadass Sudhir provides a marginal overview for future research and improvements.

III. SYSTEM DESCRIPTION

Graphical representation of any process is always better and more meaningful than its representation in words. A euclidean distance can be efficiently used to identification of acute appendicitis. This system consists of five steps,

(a) The first step is image capturing; here the ultrasound image is acquired from patient database through mat lab

(b) The second step is image preprocessing; remove of noise components appendicitis regions such as label, marks on the image is removed

(c) The third step is image enhancement; noise in the preprocessed image is removed through median filter

(d) The fourth step is used similarity measures include euclidean and manhattan distance;

(e) The final step is matched the results to manhattan and Euclidean distance; here to find out the acute appendicitis.

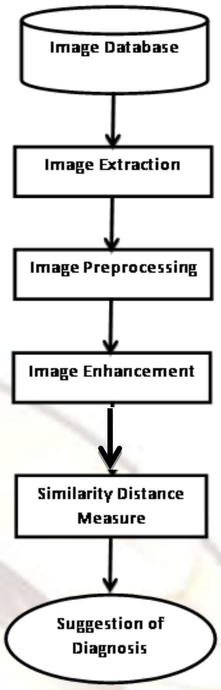


Fig3.1. Overview of proposed system

IV. PROPOSED ALGORITHM

To summarize our method authors described below, the algorithm is as follows.

Step1: Image Capturing

Image Capturing is a process to acquire the digital image into Matlab. Ultrasound image from the database is obtained through imread command in matlab. The exact path of the image should be given as the argument to imread command. In matlab, captured image is displayed using imshow command by passing a variable as the argument and the image can also be displayed in the image viewer using imtool command.

Step2: Image Extraction

Appendicitis is slightly brighter than its surrounding areas, produces a sharp peak of unusual gray level intensity pixels.

Step3: Image Prepreocessing

The pre-processing technique eliminates the incomplete, noisy and inconsistent data from the image in the training and test phase. In the images are too noisy or blurred, they should be filtered and sharpened and for removing the unwanted portions of the image.

Step4: Image Enhancement

The preprocessed images will have some noise which should be removed for the further processing of the image. In conventional enhancement techniques such as low pass filter, median filter, gabor filter, gaussian filter, prewitt edge-finding filter, normalization method are employable for this work.

Step5: Similarity Measures

In this paper, the appendicitis is detected using the distance measure in order to confirm the patient is diagnosed with appendices. The distance is predominant to bring out the diagnosis [19].

$$\textbf{Euclidean:}\ dis(t_i, t_j) = \sqrt{\sum_{h=1}^{k} (t_{ih} - t_{jh})^2}$$
$$\textbf{Manhattan:}\ dis(t_i, t_j) = \sum_{h=1}^{k} |\ (t_{ih} - t_{jh})\ |$$

Step6: Suggestion of Diagnosis

V. EUCLIDEAN DISTANCE

Euclidean distance is the distance between two points in Euclidean space. Take two points P and Q in two dimensional Euclidean spaces. This describes P with the coordinates (p1, p2) and Q with the coordinates (q1, q2). Now construct a line segment with the endpoints of P and Q. This line segment will form the hypotenuse of a right angled triangle. The distance between two points p and q is defined as the square root of the sum of the squares of the differences between the corresponding coordinates of the points; for example, in two-dimensional Euclidean geometry, the Euclidean distance between two points a = (ax, ay) and b = (bx, by) is defined as:

$$d(a, b) = \sqrt{(p_1 - q_1)^2 + (p_2 - p_2)^2}$$

This algorithm computes the minimum Euclidean distance between a column vector x and a collection of column vectors in the codebook matrix cb. The algorithm computes the minimum distance to x and finds the column vector in cb that is closest to x. It outputs this column vector, y, its index, idx, cb, and distance, the distance between x and y.

$$d(a, b) = |p - q|$$

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2}$$

$$= \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}$$

In one dimension, the distance between two points, x1 and x2, on a line is simply the absolute value of the difference between the two points: [3]

$$\sqrt{(X_2 - X_1)^2} = |X_2 - X_1|$$

In two dimensions, the distance between P = (p1,p2) and
q = (q1,q2) is:

$$\sqrt{(p_1 - q_1)^2 + (p_2 - p_2)^2}$$

VI. MANHATTAN DISTANCE

The distance between two points in a grid is based on a strictly horizontal and/or vertical path as opposed to the diagonal. The Manhattan distance is the simple sum of the horizontal and vertical components, whereas the diagonal distance might be computed by applying the Pythagorean Theorem [16].

The Manhattan distance function computes the distance that would be traveled to get from one data point to the other if a grid-like path is followed and Manhattan distance between two items is the sum of the differences of their corresponding components. Fig 5.1 shows the difference between Euclidean Distance and Manhattan Distance. The formula for this distance between a point X=(X1, X2, etc.) and a point Y=(Y1, Y2, etc.) is [15]:

$$d = \sum_{i=1}^{n} |x_i - y_i|$$

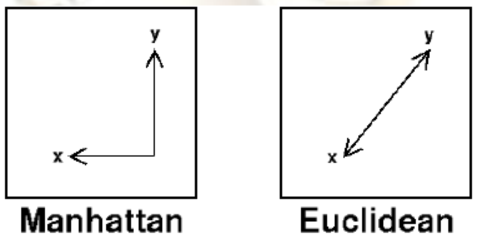


Fig. 6.1 Difference between Euclidean Distance and Manhattan Distance [32]

VII. EXPERIMENTS AND RESULTS

An experiment has been conducted on sonographic scan image based on the proposed flow diagram as shown in Fig 3. 1. The sonographic image is first captured as an input image, and then decorrelation stretch has been used to remove the noise component from that image. Decorrelation stretch has been used to find the edge feature in the ultrasound scan image. Using this technique the noise component is removed from the image.

7.1 Sample Appendicitis

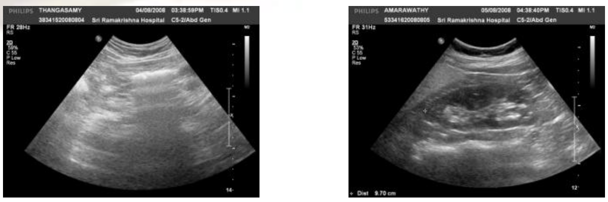


Fig. 7.1 Normal Image vs Appendicitis images

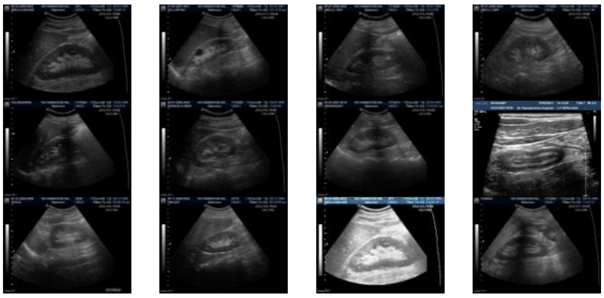


Fig. 7.2 Sample Appendicitis images

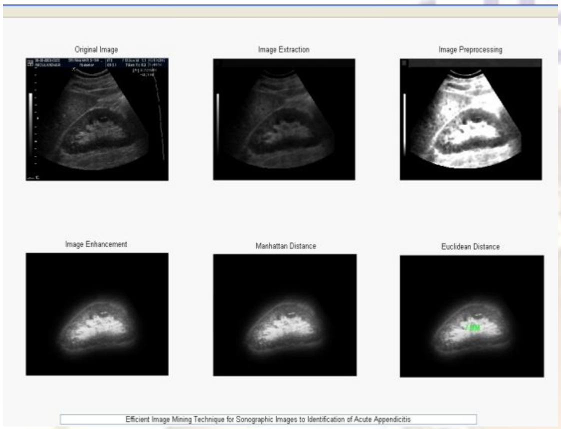
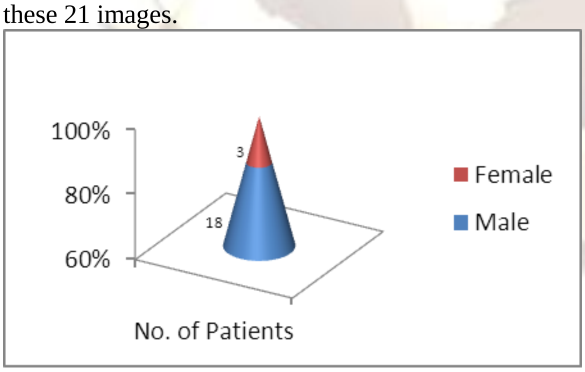


Fig. 7.3 Proposed Systems shows Appendicitis images

7.2 Results

The study period was from Aug 2011 – Nov 2011. Ultrasound imaging was done in all patients. Patients were followed up until the discharge diagnosis was made. Acute appendicitis of sonographic image findings in 21 patients was done. In the experimental study the total number of 21 instances studied where in the range of 16 – 42 years. Male and female sex ratio is in the range of 6: 1. The column chart clearly shows the sex distribution for these 21 images.



Graph 7.1 Patients Ratio

The images are classified in two different sizes based on the thickness of appendicitis with greater than 6 mm and less than 6 mm. The proposed system is tested with 21 images and the results obtained are: out of 21 images, 21 instances show thickness measured as greater than 6 mm.

Table 6.1 : Comparison of Proposed Method, Real Result, WBC, Xray and Manhattan Distance

S NO	AGE	SEX	Real Result	WBC	X RAY	Eculidean	Manhattan
1	42	M	+	+	-	+	-
2	26	M	+	+	-	+	-
3	31	M	+	+	-	+	-
4	25	M	+	+	-	+	-
5	18	M	+	+	-	+	-
6	21	M	+	+	-	+	-
7	40	M	+	+	-	+	-
8	21	M	+	+	-	+	-
9	24	M	+	-	-	+	-
10	29	M	-	+	-	+	-
11	34	M	+	+	-	+	-
12	24	M	+	+	-	+	-
13	27	M	-	+	-	+	-
14	42	M	+	+	-	+	-
15	35	M	+	+	-	+	-
16	37	M	+	-	-	+	-
17	40	M	+	-	-	+	-
18	16	M	+	+	-	+	-
19	22	F	+	+	-	+	-
20	28	F	+	-	-	+	-
21	30	F	+	-	-	+	-

A sonologist has been consultated and the results obtained from the sonologist on the same 21 samples reveal that 21 patients are affected by appendicitis. A comparison of the results obtained from the proposed system with the results obtained from the sonologist has been done and the results are as shown in table 6.2.

VIII. Performance Evaluation

All these features were extracted for this particular domain. Authors used the features extracted from the training image, on the testing images. Hence in essence, authors have developed a system, which is trained once and then applies the same technique to other images. The goodness of these features can be judged by certain evaluation criteria. The second set of images was used for comparison and validation. The validation images have the same size and hence can be compared with the extracted images. For a two-class problem, there can be 4 possible outcomes of a prediction [17]. The outcomes are True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN). Based on commonly used performance measures in two statistical measures were computed to assess system performance namely recall and precision. Table8.1 shows the values of recall and precision of each diagnosed images using the proposed system.

Precision: Defined as the diagnosed images, which is relevant.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Recall: Defined as the diagnosed images versus all database images.

$$\text{Recall} = \frac{TP}{TP + FN}$$

Table 8.1 Recall and precision using the proposed system

Method	Precision	Recall
Eculidean	0	1
Manhattan	0.5	1
Real Result	0.525	1
WBC	0.567	1
XRay	0.5	1

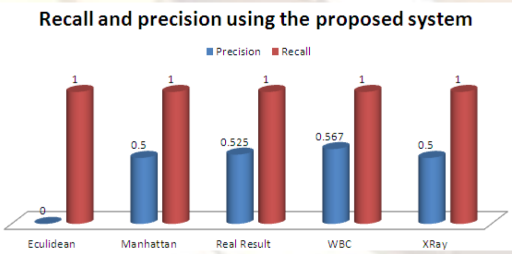


Fig. 8.1 Recall and precision using the proposed system

XI. CONCLUSION AND FUTURE WORK

In this paper, authors propose an automatic appendicitis detection system that takes sonographic images as input and using image mining technique. The images are preprocessed with various techniques. Compared with real result, WBC, Xray and Manhattan, the proposed system offers many advantages including better accuracy, greater noise reduction, faster speed and complete automation. The precision and recall obtained are 0.525 and 1 respectively. The developed system is expected to provide valuable support to the the physicians in decision making at the time of diagnosis. *This work can be further extended to improve the precision of the diagnosis result while detecting appendicitis from sonographic images.*

ACKNOWLEDGMENT

Authors wish to thank Sonologist for providing the sonography images. The authors extended their sincere thanks to Department of Computer Applications for giving permission to do the research and all others directly or indirectly helped in pursuing this research.

REFERENCES

[1] S. G. Prabhudesai S. Gould S. Rekhraj P. P. Tekkis G. Glazer P. Ziprin, "Artificial Neural Networks: Useful Aid in Diagnosing Acute Appendicitis", *World J Surg* (2008) 32:305–309.

[2] E.Sivasankar, S.Rajesh andR.Venkateswaran,"Diagnosing Appendicitis Using Backpropagation Neural Network and Bayesian Based Classifier", *International Journal of Computer Theory and Engineering*, Vol. 1, No. 4, October, 2009 1793-8201

[3] R. Balu, T. Devi, "Identification of Acute Appendicitis Using Euclidean Distance on Sonographic Image", *International Journal of Innovative Technology & Creative Engineering*, ISSN : 2045-8711, pp 32 - 38, Vol.1 No.7 July 2011

[4] Mesut Tez Selda Tez Erdal Gocmen, "Neurofuzzy is Useful Aid in Diagnosing Acute Appendicitis", *World J Surg* (2008) 32:2126.

[5] Ji Zhang, Wynne Hsu and Mong Li Lee, "Image Mining: Trends and Developments", *Journal of Intelligent Information Systems archive*, Volume 19, Issue 1 Pages: 7 - 23, 2002.

[6] Jose M. Merigo, "Decision making with distance measures, owa operators and weighted averages", gandalf.fcee.urv.es/ sigef/english/congressos/.../045_Merigo.pdf

[7] Sandy Craig and Barry E Brenner, "Appendicitis",http://emedicine.medscape.com/articl/773895-overview

[8] Zhang Ji, Hsu, Mong and Lee, "Image Mining: Issues, Frameworks and Techniques," *Proceedings of the Second International Workshop on Multimedia Data Mining (MDM/KDD'2001)*, in conjunction with ACM SIGKDD conference, San Francisco, USA, 26 August.

[9] C. Ordonez and E. Omiecinski (1999), "Discovering association rules based on image content," *Proceedings of the IEEE Advances in Digital Libraries Conference* (ADL'99).

[10] R. Missaoui and R. M. Palenichka (2005), "Effective image and video mining: an overview of modelbased approaches," In MDM '05: *Proceedings of the 6th international workshop on Multimedia data mining*,

pp. 43–52.

[11] Ashok N. Srivastava and Nikunj Oza (2004), "Knowledge Driven Image Mining With Mixture Density Mercer Kernels," *Workshop on the Theory and Applications of Knowledge Driven Image Information Mining,* with focus on Earth Observation.

[12] P. Rajendran and M. Madheswaran (2009), "An Improved Image Mining Technique For Brain Tumour Classification Using Efficient classifier", *International Journal of Computer Science and Information Security*, vol. 6, no. 3, pp. 107-116.

[13] Rajshree S. Dubey (2010), "Image Mining using Content Based Image Retrieval System", (IJCSE) *International Journal on Computer Science and Engineering* Vol. 02, No. 07, pp. 2353-2356.

[14] Ramadass Sudhir, "A Survey on Image Mining Techniques: Theory and Applications", *Computer Engineering and Intelligent Systems* , ISSN 2222-1719, Vol 2, No.6, 2011

[15] Predictive Patterns Software Inc. in January of 2004 "Overview of Manhattan Distance" at http://www.improvedoutcomes.com/docs/WebSiteDocs/Clustering/Clustering_Parameters /Manhattan_Distance_Metric.htm

[16] Wikipedia, Mar 25, 2012. "Manhattan distance", http://en.wiktio nary .org/ wiki/Manhattan_distance,

[17] I. Witten and E. Frank, "Data Mining: Practical Machine Learning tools and techniques with Java Implementations", *Morgan Kaufmann publications*, 2000.

[18] Fatma, S.N., Nashipudimath, M, "Image mining using association rule", *Information and Communication Technologies* (WICT), 2011.

[19] Margaret H. Dunham, "Data Mining: Introductory and Advanced Topics", *Prentice Hall*, 2002