

Developing an Intelligent Decision Support System for the Diagnosis of Some Children's Diseases

A. E. E. Elalfi
Computer science
Department
Faculty of Specific Education
Mansoura University

M. A-H. Fouda
Heart and Chest Surgery
Department
Faculty of Medicine
Mansoura University

A. A. Atta
Computer Science
Department
Faculty of Specific
Education
Mansoura University

ABSTRACT

The aim of this paper is to present an intelligent system for the diagnosis of some children's diseases to help fresh and inexperienced healthcare graduates. This system is based on clinical database, knowledge base and medical image processing. This intelligent system provides a graphical user interface which allows the user to choose among a number of symptoms and input a medical diagnostic image to get the accurate diagnosis.

Keywords

Artificial intelligent, Knowledge base, Image database, Intelligent Systems, Children's diseases

1. INTRODUCTION

The advancement in computer technology has empowered the software developers and domain knowledge experts to build more intelligent tools for assisting medical practitioners in making their decisions [1]. In medicine the relationship between disease and symptoms is hardly ever one to one, so the range of symptoms is inherently difficult for novice doctor. An intelligent system can resolve real world problems using human knowledge and following human reasoning skills [2].

Expert systems or knowledge-based systems are the commonest type of Artificial Intelligent (AI) systems in routine clinical use. They contain medical knowledge, usually about a very specifically defined task, and are able to reason with data from individual patients to come up with reasoned conclusions. Although there are many variations, the knowledge within an expert system is typically represented in the form of a set of rules [3].

Chest diseases are one of the greatest health problems for people living in the developing world. Millions of people are diagnosed every year with a chest disease in the world [4].

Recurrent chest infections are a common reason for children to be seen by their General Practitioner (GP) or a pediatrician. Persistent cough may be the only symptom, but often there is also a history of wheeze, breathlessness, sputum production or general ill-health [5].

Children often have acute, short-term illnesses such as upper respiratory tract or ear infections, gastrointestinal illness with vomiting and diarrhea, or injury-related problems. However, some children develop chronic illness (lasting for years or even lifelong) as a result of Jain et al. [10] have presented a computer aided method for the detection of Melanoma Skin Cancer using Image

genetic (inherited) conditions, or a combination of both [6].

The challenge for the clinician is to distinguish between the child with self-limiting or minor problems and the child with serious, perhaps progressive, lung disease. To do this, it is necessary to assess the severity and to diagnose the cause of the symptoms [5].

The paper is organized as follows: section 1 includes introduction, section 2 contains related work, section 3 presents knowledge acquisition, section 4 presents knowledge representation, section 5 implicates the advisory system, section 6 contains application and results, and section 7 is conclusions and future work.

2. RELATED WORK

Computer-aided diagnosis (CAD) has been developed fast in the last two decades. Some pioneer studies are dated back to the 1960s. In 1998, the first U.S. Food and Drug Administration (FDA) approved commercial CAD system [6].

Samy et al. [7] have presented an expert system aims to provide the patient with background for suitable diagnosis of some of the eye diseases by using CLIPS language as a tool for designing their expert system. This expert system uses the following categories (1), Rule-based systems (2), Knowledge-based systems (3), Intelligent agent (4), Database (5), methodology (6), Inference engine (7), System-user interaction.

Nusai et al. [8] have represented an automated pilot system to assist in diagnosis of metabolic brain diseases, they use the diffusion-weighted imaging which are preprocessed to produce exponential apparent diffusion coefficient images which are later brain extracted and normalized to a standard brain template. They utilize wavelets to clear these images. The images corrected, threshold and then the abnormal regions are identified using different brain atlases. Then a fuzzy relational classifier using the abnormal region features to categorize the disease.

Tenório et al. [9] have introduced a web-based system for diagnosing celiac disease which constructs a retrospective database that included 178 clinical cases for training. Tests were conducted on 270 automated classifiers by using five artificial intelligence techniques, namely decision trees, Bayesian inference, k-nearest neighbor algorithm, support vector machines and artificial neural networks, the parameters evaluated were accuracy and sensitivity.

Processing tools, the input to the system is the skin lesion image and then by applying image processing

techniques to conclude about the presence of skin cancer. The Lesion Image analysis tools check for the various Melanoma parameters like Asymmetry, Border, Color and Diameter by texture, size and shape analysis for image segmentation and feature stages. The extracted feature parameters are used to classify the image as Normal skin and Melanoma cancer lesion.

Keles. A et al. [11] have developed an expert system for diagnosis of breast cancer, because difference between benign and malignant mammographic findings, however, is quite difficult. Only 15–30% of biopsies performed on no palpable but mammography suspicious lesions prove malignant, the golden standard for diagnosis of breast cancer is biopsy but biopsy can be a source of patient discomfort, bleeding and infection. It can burden the health care system with extra costs. Thus, to reduce unnecessary biopsy rate has acquired great importance.

This paper motivates to serve both the doctors and patients in providing decision support system. Therefore, it presents an advisory system for the diagnosis of some children's chest diseases based on medical knowledge database and image processing of medical images. The structure of the proposed system is presented in the following sections.

3. CHILDREN'S CHEST DISEASES KNOWLEDGE BASE

In the current study medical concepts and knowledge are acquired from several sources:

3.1 Domain Knowledge Base

The medical background of children's chest diseases is recorded through the creation of personal interview with doctors.

Table1 shows sample knowledge which is used in proposed system.

Concept, property and values are current used tools to acquire knowledge. Each concept has two or three properties.

3.2 Medical Images Knowledge Base

A set of medical images associated with children's chest diseases have been collected from hospitals, radiological centers, medical multimedia and the internet, these images have been subjected to a processing operations where its special features have been extracted and saved in a data base and in case of entering a new image (Query image) the special features of this image are extracted and compared with the previously saved special features of the images in the data base. From the comparison, 90 images which are most similar to the query image can be seen; the decision will be taken based on the first one of them. The percentage of similarity resulting from this comparison supports the diagnosis decision.

A set of medical images is shown in figure 1.

Table 1. Sample of acquired diagnosis knowledge

Concept	Property		Value	Abbr.
Head	State	H_S	Headache	H_D
	temperature	H_T	High	H_G
Throat	State	TH_S	Inflammation	IN_F
			Normal	N
			Throat Bain	TH_B
	Congestion	C_G		
Color	TH_C	Redness	R_D	
Mouth	State	M_S	Difficulty in swallowing	D_S
			Dry and chapped lips	D&C_L
			Breathing from mouth	B_M
Color	M_C	Blue lip	B_L	
Tonsils	State	T_S	Enlarged tonsils Or Pus spots on tonsils	E_T OR P_S_T
Nose	State	N_S	Dilated nostrils	D_N
			congestion	C_G
			Rhinitis or snoring or stuffy nose	RH Or SN Or ST_N
Eyes	Color	E_C	Redness	RED
Chest	State	CH_S	Increase in breathing rate Or Wheezing Or cough Or chest contractions	IN_BR_ Rate Or WH Or CO Or CH_ N
			Increase in breathing or rate Wheezing or Sever cough or contractions	IN_BR Or R_WH Or S_CO Or CON
			contractions persistent cough or or difficulty in breathing or wheezing or increase in mucus secretion	CON_P Or CO Or D_BR Or WH Or IN_M_S ECTION
			contractions persistent or cough sever or difficulty in breathing or wheezing or increase in mucus secretion	CON_P Or S_CO Or D_BR Or WH Or IN_M_S ER
			Difficulty in breathing or Dry cough	D_BR Or D_CO

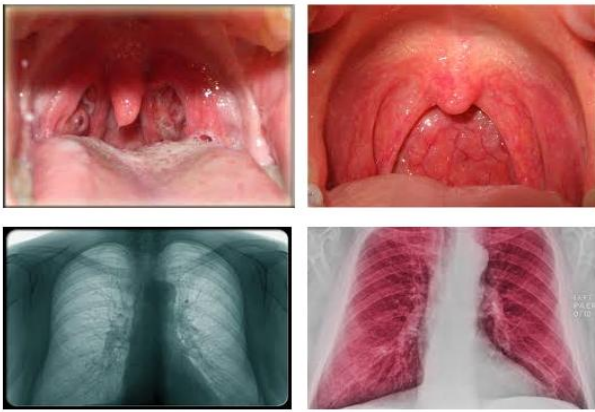


Fig 1: An example for sore throat and chest infection images

3.2.1 Image Processing

Diagnosis of diseases in this study was based on clinical data and medical images.

Image processing is the field of research concerned with the development of computer algorithms working on digitized images [12].

Medical diagnosis is one of the most important area in which image processing procedures are usefully applied. Image processing is an important phase in order to improve the accuracy for diagnosis procedure [13].

Image processing in this study passed through several stages:

A) Image acquisition

In this study images are collected through experts and internet. We select eight different classes of children's diseases: (Sore throat, Tonsillitis, Bronchitis, Severe bronchitis, Pneumonia, Whooping cough, Asthma, Kawasaki disease).

B) Image pre-processing

The input image given to the system can be obtained in noise or corruption hence it needs to be pre-processed. Here, the pre-processing includes the image resizing and enhancing it from noising and corruption, in addition to convert image to grey.

C) Features Extraction

All features extracted in this study are based on texture analysis using Grey-Level Co-occurrence Matrix (GLCM).

The GLCM is a powerful tool for image feature extraction by mapping the grey level co-occurrence probabilities based on spatial relations of pixels in different angular directions [14].

The main features that extracted from GLCM are [15].

• Contrast

Contrast is a local grey level variation in the grey level co-occurrence matrix. It can be thought of as a linear dependency of grey levels of neighboring pixels.

$$\text{Contrast} = \sum_{I,J} |I - J|^2 P(I, J) \quad (1)$$

Where, I and J are the horizontal and vertical cell coordinates and p is the cell value. If the neighboring pixels are very similar in their grey level values, then the contrast in the image is very low. In case of texture,

the grey level variations show the variation of texture itself. High contrast values are expected for heavy textures and low for smooth, soft textures. The range of Contrast is $[0, (\text{size}(\text{GLCM}, 1) - 1)^2]$ where Contrast is 0 for a constant image.

• Homogeneity

Homogeneity measures the uniformity of the non-zero entries in the GLCM. It weights values by the inverse of contrast weight.

$$\text{Homogeneity} = \sum_{I, J} \frac{1}{1 - (I - J)^2} P(I, J) \quad (2)$$

The GLCM homogeneity of any texture is high if GLCM concentrates along the diagonal, meaning that there are a lot of pixels with the same or very similar grey level value. The larger the changes in grey values, the lower the GLCM homogeneity making higher the GLCM contrast. The range of homogeneity is $[0, 1]$. If the image has little variation then homogeneity is high and if there is no variation then homogeneity is equal to 1. Therefore, high homogeneity refers to textures that contain ideal repetitive structures, while low homogeneity refers to big variation in both, texture elements and their spatial arrangements. An inhomogeneous texture refers to an image that has almost no repetition of texture elements and spatial similarity in it is absent.

• Energy

Energy is a measure of local homogeneity and therefore it represents the opposite of the Entropy. Basically this feature will tell us how uniform the texture is.

$$\text{Energy} = \sum_{i,j} p(i, j)^2 \quad (3)$$

• Correlation

$$\text{Correlation} = \frac{\sum_{I=0}^{G-1} \sum_{J=0}^{G-1} ijP(i, j) - (m_i m_j)}{\sigma_i \sigma_j} \quad (4)$$

Where, i and j : row and column numbers in the GLCM matrix, σ_i and m_i are the mean and standard deviation of P (i, j) rows, m_j and σ_j the mean and standard deviation of P(i, j) columns, respectively [16].

3.2.2 Classification

Images were classified by using the Weighted Euclidean Distance measure (WED), it is one of the fundamental and widely used techniques in matching. It is the technique used in the proposed system for bacterial image matching. The formula of WED measure can be written as follows [17].

$$d(v, v^k) = \sqrt{\sum_{i=1}^n p_i (v_i - v_i^k)^2} \quad (5)$$

Where, v_i to balance the variations in the dynamic range.

p_i the weight added to the component.

k is the matched image index.

$$p_i = \frac{N}{\sum_{k=1}^N (v_i^k - \bar{v}_i)^2} \quad (6)$$

N = the number of images in databases.

$$\bar{v}_i = \frac{\sum_{k=1}^N v_i^k}{N} \quad (7)$$

GUI for image classification as shown in figure 2 .

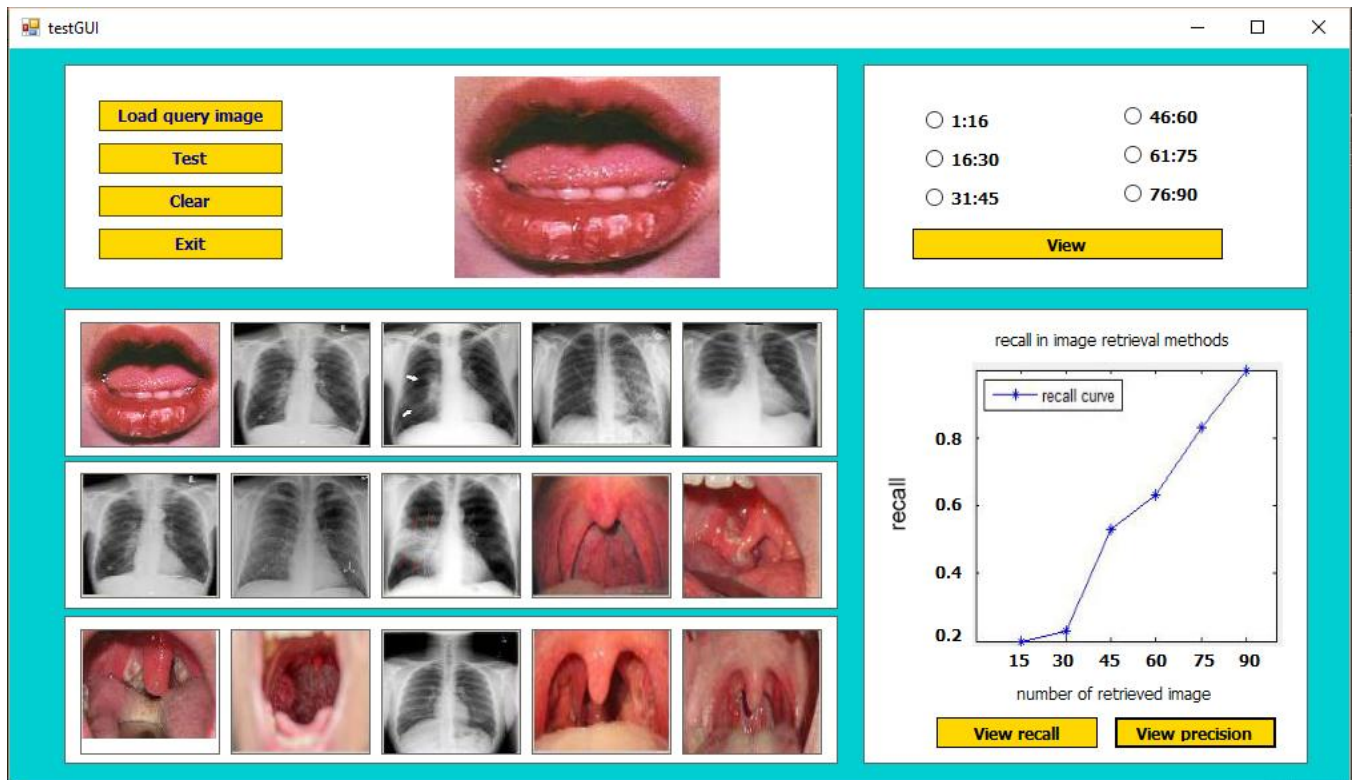


Fig 2: The graphical user interface of the image classification.

3.3 Knowledge Representation

knowledge base (KB) is the body of subject knowledge that supports the performance of a "knowledge-based system," such as an expert or natural language processing system [18].

For representing knowledge there are many methods such as Semantic Networks, Frames, Production Rules and Java Classes [19].

Production rules are the most used method for representing knowledge because it's a plausible model of human problems solving.

Sample of production rule for the diagnosis of children's chest diseases are shown as follows:

Rule1:

IF H_S=H_D ^ H_T=H_G ^ TH_S=IN_F ^ TH_C=Red ^ M_S=D_S ^ Image = Sore Throat THEN

Disease = Sore Throat 100%

Rule2:

IF H_S=H_D ^ H_T=H_G ^ TH_S=T_B ^ M.S=D & CH_L ^

T_S= E_T OR P_S_T ^ N_S=C_G ^

Image = Tonsillitis THEN

Disease = Tonsillitis 100%

The abbreviation which used are explained in table 1.

The final diagnosis decision resulting from a union between the diagnosis percentage resulting from the knowledge base and the percentage of similarity resulting from images processing.

4. ADVISORY SYSTEM

The advisory system diagnosis children's chest diseases by using KB and medical images (image processing) if they are exist.

The KB includes knowledge about (Sore throat, Tonsillitis, Bronchitis, Severe bronchitis, Pneumonia, whooping cough, Asthma, and Kawasaki disease), the final diagnosis displays the proposed chest diseases with the probability.

The flowchart for children's chest diseases diagram is shown in figure 3.

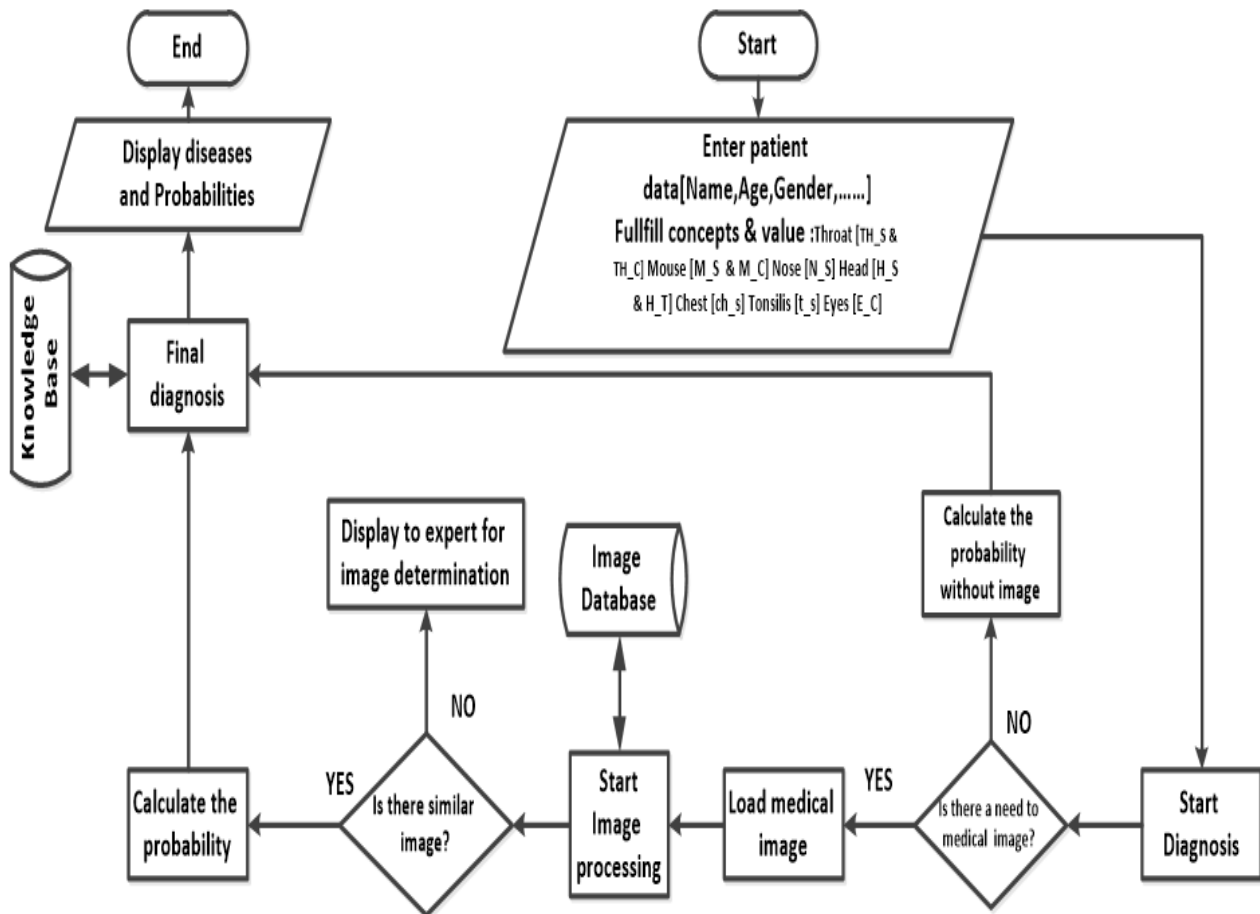


Fig 3: The flowchart for children’s chest diseases

Vin Diagram which represent the decision probability for diagnosis children's chest diseases is shown in figure 4.

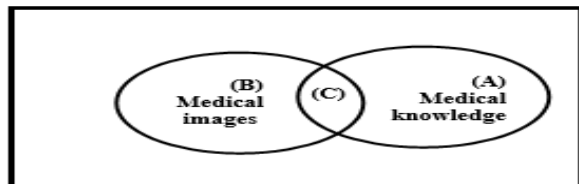


Fig 4: Vin diagram for decision probability

A → Refers to the decision probability for properties of main concepts only without image.

B → Refers to the decision probability for image only.

C → Refers to the decision final probability for properties of main concepts and image.

• **The Graphical User Interface**

Communication between the user and the system is done through the Graphical User Interface (GUI) [5]. The user interface is represented as a menu which displays the symptoms diseases to the user. When the system is started, the main menu is displayed on the screen which asks the user to choose specific property value of any main concept that fits its condition.

GUI on this study consists of a main menu divided into 7 tabs, every tab is specific for a concept. Each concept has two or three properties, each property has a value or a number of values which allow the user to choose between symptoms and input medical images that could support the diagnosis.

The GUI of the proposed system as shown in figure 5.

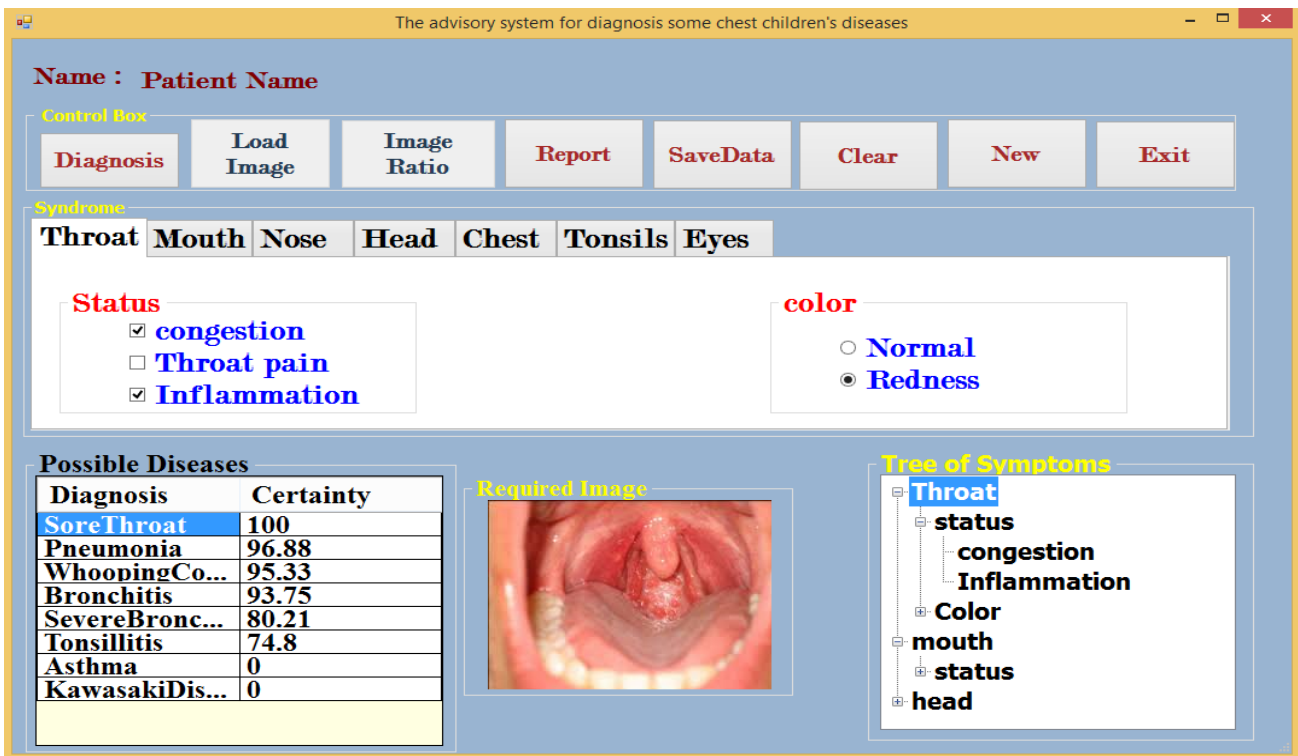


Fig 5: The graphical user interface of the advisory system

5. APPLICATIONS & RESULTS

The suggested system was applied to two groups, each of them have 50 cases. The system was found matched with a number of cases and did not match with others. Table 2 shows the number of cases which the system matches with and the number of cases it didn't match with in each group. Also it shows what was updated. So, the knowledge base was updated according to the unmatched cases.

Table 2. The percentage of agreement and disagreement among children's diseases experts

	Num. of cases	Agree	Disagree	%	Action
First Group	50	45	5	90%	Adding one symptom for tonsillitis (Fever)
Second Group	50	47	3	94%	Adding a new disease (Gastroenteritis)
Total	100	92	8	92%	

The performance of the system images retrieval can be measured by using Recall and Precision. Recall measures the ability of the system to retrieve all the models that are relevant, while Precision measures the ability of the system to retrieve only the models that are relevant [20].

$$\text{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}} \quad (8)$$

$$\text{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images in database}} \quad (9)$$

Figure 7 and figure 8 show recall and precision curve for content based chest diseases image retrieval methods.

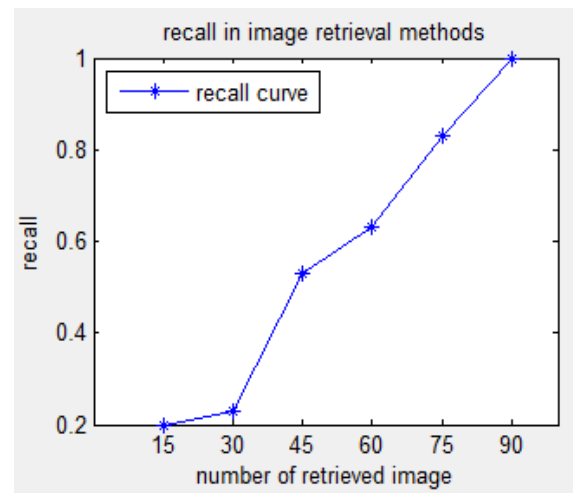


Fig 7: The recall curve

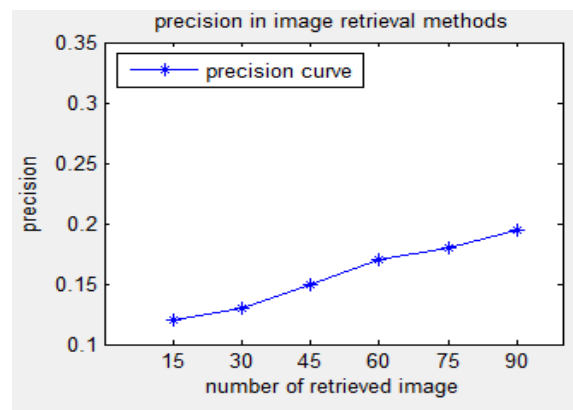


Fig 8: The precision curve

6. CONCLUSIONS & FUTURE WORK

System application can't be done unless a high degree of trust was reached. The cases record can be useful in increasing the system learning. The system is useful in diagnosis decision by helping the decision makers. The new techniques are important in advisory systems especially in diseases diagnosis such as image processing and sound processing.

The results can be improved in future work by applying the proposed system on different platforms such as mobiles and tablets, extending the advisory system of this study to other fields and categories and integrating image processing, sound processing, and knowledge base in building a diagnostic advisory system.

7. ACKNOWLEDGMENTS

The authors are indebted to Dr. Mohamed ALshal, pediatrician, and Dr. Hossam AL sayed, pharmacist, for their participations in this work.

8. REFERENCES

- [1] Dutta, T., et.al., 2015. An Intuition Based Fuzzy Logic Driven Approach for Designing Symptomatic Medical Diagnostic Expert System, International Journal of Engineering Research & Technology (IJERT), 4, 876-878.
- [2] Prasatl, P., Prasad, K., and Sagar, Y., 2011. AN Approach to develop expert system in medical diagnosis uses machine learning algorithm (ASTHMA) and performance study, International Journal on Soft Computing (IJSC), 2, 26-33.
- [3] Orhan, E., Yumusak, N., and Temurtas, F., 2012. Diagnosis of chest diseases using artificial immune system, Expert Systems with Applications, 39, 1862-1868
- [4] Fujita, H., Uchiyama, Y., and Nakagawa, T., 2008. Computer-aided diagnosis: The emerging of three CAD systems induced by Japanese health care needs. computer methods and programs in biomedicine, 92, 238-248.
- [5] Jon Couriel., 2002. Assessment of the child with recurrent chest infections. British Medical Bulletin.61, 116-132.
- [6] Janet M. Torpy, MD, 2010. Chronic Diseases of Children. The Journal of the American Medical Association (JAMA), 303, 7.
- [7] Samy, S., and Abu-Zaiter, A., 2008. An expert system for diagnosis eye diseases using clips. Journal of Theoretical and Applied Information Technology.
- [8] Amir A. Sepehri, Joel Hancq, Thierry Dutoit, Arash Gharehbaghi, Armen Kocharian and A. Kiani, "Computerized screening of children congenital heart diseases", Computer Methods and Programs in Biomedicine, Vol.92, Issue 2, November 2008, PP. 186-192.
- [9] Tenório, J., Hummel, A., and Cohrs, F., 2011. Artificial intelligence techniques applied to the development of a decision-support system for diagnosing celiac disease.

International journal of medical informatics. 80, 763-802.

- [10] Jain, S., Jagtap, V., and Pise, N., 2015. Computer aided Melanoma skin cancer detection using Image Processing. International journal of Computer Science. 48,735-740.
- [11] Kekes, A., Ayturk, A., and Yavuz, U., 2011. Expert system based on neuro-fuzzy rules for diagnosis breast cancer. Expert system with Application. 38, 5719-5726.
- [12] De Ridder, D., et.al., Nonlinear image processing using artificial neural networks.
- [13] Arena, P., Basile, M., Bucolo, M., and Fortuna, L., 2003. Image processing for medical diagnosis using CNN. International journal of Computer Science.479, 174-178.
- [14] Jadhav, S., and Kamat D., 2014 Analysis and Detection of Skin Cancer. Journal of Electronics and Communication Engineering.9, 50-54.
- [15] Gebejes, A., and Huertas, R.,2013. Texture Characterization based on Grey-Level Co-occurrence Matrix. Journal of Informatics and Management Sciences. 12, 375-378.
- [16] Vargas, J., et.al., 2011. offline signature verification based on gray level information using texture features", Pattern Recognition, 44, 375-385.
- [17] R. Mukundan and K. R. Ramakrishnan., Moment Functions In Image Analysis Theory and Applications, World Scientific Publishing Co. Pte. Ltd.1998 ISBN 981-02-3524-0 , 81-85.
- [18] Travis, L., 1990. Knowledge Representation in Artificial Intelligence. Artificial intelligence and expert systems, 215-238.
- [19] John, F., 2000. Knowledge Representation: Logical, Philosophical, and Computational Foundations, Computational Linguistics.27, 287-294.
- [20] Danish, M., Rawat, R., and Sharma, R., 2013. A Survey: Content Based Image Retrieval Based On Color, Texture, Shape & Neuro Fuzzy, International Journal of Engineering Research and Applications.3, 839-844

9. AUTHOR'S PROFILE

Dr. A.E. El-ALfy, Professor Faculty of Specific Education, Computer department, Mansoura University. Scientific degrees: In 1973 he earned a bachelor's degree with a very good grade, then in 1986 he earned a master's degree with honors and in 1992 he earned his doctorate with honors. He became professor in 2007.

Dr. M. A-H. FOUDA, Professor Emeritus of Chest and Heart Surgery, Faculty of Medicine, Mansoura University.

A. A. Atta, Demonstrator, Department of Computer Science Faculty of Specific Education, Mansoura University, Cairo, Egypt, She got her Bachelor degree in 2011, and she is M.Sc.student.