ISSN 2321-9017
Volume 1, No.1, June – July 2013
International Journal of Bio-Medical Informatics and e-Health
Available Online at http://warse.org/pdfs/ijbmieh03112013.pdf

A Hybrid Case-Based and Content-Based Retrieval Engine for Mobile Cancer Management System - MCMS

Bassant Mohamed ElBagoury, Abdel-Badeeh M.Salem, Mohamed Ismail Roushdy

Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt

bassantai@yahoo.com , abmsalem@yahoo.com, miroushdy@hotmail.com

Abstract: **-** In this paper, we introduce a Case-Based Reasoning ***(CBR)*** Engine prototype for Mobile Remote Diagnosis of cancer patients. Moreover, the retrieval in ***CBR*** is a very difficult complex task for medical diagnosis. This is due to diagnostic radiology that requires accurate interpretation of complex tumour features in medical images. This may lead to uncertain diagnosis decisions. In this paper, we also propose a new hybrid retrieval algorithm for breast cancer diagnosis. It combines Content-Based Image retrieval ***CBIR*** techniques that could be valuable to radiologists in assessing medical images by identifying similar images that could assist with decision support. The proposed retrieval model is applied to FNA breast cancer images and it is tested with 600 radiology images. Cross-validation test has shown an average retrieval accuracy of 90%. The hybrid model is described in the context of a prototype namely, MCMS a hybrid expert system, which integrates CBIR into the CBR.

Key-Words: - Mobile Cancer Management System - MCMS, Artificial Intelligence, Content-Based Image Retrieval, Teleradiology, Case-Based Reasoning.

1 Introduction

Mobile teleradiology is a steadily growing field in telemedicine, and it combines recent developments in mobile communications and network technologies with telemedicine applications [14, 15]. However, the main problem still remains in providing intelligent remote medical experts consultations for emergency cases.

However, those mobile remote diagnosis systems need access to huge sized medical image databases at hospitals servers. These have posed technical challenges to computer systems to store/transmit and index/manage image data effectively to make such large collections easily accessible. Storage and transmission challenges are taken care by Image compression. The challenges of image indexing are studied in the context of image database [1,3,5], which has become one of the most important And promising research area for researchers from a wide range of disciplines like computer vision, image processing and database areas. The need for faster and better image retrieval techniques is increasing day by day. One of the most important applications is for medical imaging is Content-Based Image Retrieval [CBIR] [4,6,7,16].

CBR has long been applied in medicine [11]. The retrieval accuracies are not accurate. Despite of the importance of the retrieval task, most CBR medical systems lack the implementation of retrieving radiology images, which are crucial to patients' diagnosis[. Correspondingly, many techniques have been developed for fast indexing and retrieval of digital images, such as Content-Based Image Retrieval [CBIR]. In this paper, we present a new hybrid case-based retrieval model for diagnosis of breast cancer patients. It combines CBIR [5,6,7,8] and CBR retrieval methods. The model is described in the context of our Mobile Cancer Management System **[MCMS]** prototype. Our motivation in this work is to extend the retrieval model for our developed prototype for computer-aided breast cancer diagnosis.

2 System Methodology

The problems of cancer treatment are extremely diverse beginning from the right diagnostic entry point for an efficient diagnostic process and ending by the documentation of cancer treatment for evaluation but also for scientific work. In the complete process there is a core problem: data management and usage of data for further evaluation such as support of diagnostic process using technologies offered by artificial intelligent methodology or quality assurance to establish evidence based diagnostic and treatment processes.

In our research work, we want to implement an Intelligent Cancer Management Service named based on a Service Oriented Architecture , which is shown to collect, manage, evaluate clinical data of cancer treatment and provide additional information and advice based on international treatment guidelines and the application of tools of artificial intelligence methodology. In a Service-oriented Architecture various services will be implemented to make data of different systems available and use tools of Case-Based-Reasoning [CBR] and Content-based Image Retrieval [CBIR]. Based on a Service Oriented Model [2] the platform can be divided into three levels (Figure 1: Basic System architecture - SOA). It consists of backend, application layer and front end. In the backend there are the main servers for the applications, database and CBR and CBIR modules. In this paper, we introduce the hybrid CBIR and CBR retrieval algorithm developed in the application layer of our system.

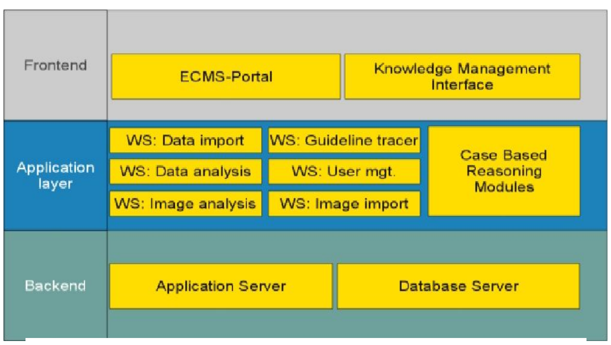

Figure1. Basic system Architecture

Content-based retrieval uses the contents of images to represent and access the images [3,4]. A typical content-based retrieval system is divided into off-line feature extraction and online image retrieval. As shown in figure 2, A conceptual framework for content-based image retrieval. it consists of four main modules, which are **Image Database**, **Feature Extraction**, **Feature Database**, **Similarity Measure.** The main algorithm of the CBIR Engine consists of four main steps, as

1. Each Image is described by its visual features such as colour, shape, texture.
2. Feature representation is represented as M-dimensional feature vector.
3. A similarity measure is used to find the similarity between a query image and database images.
4. Images are ranked in order of closeness to query and K images are returned.

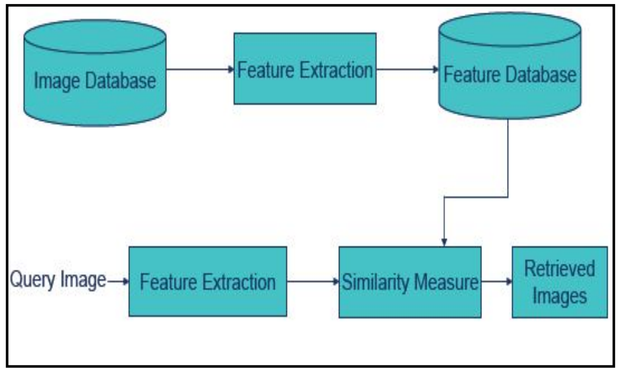

Figure 2: Basic System CBIR Retrieval Architecture

2.1 FNA Image Database

FNA Biopsy Dataset [12, 13] - Seven hundred cytology of fine needle aspiration image (i.e. cellularity, background information, cohesiveness, significant stromal component, clump thickness, nuclear membrane, bare nuclei, normal nuclei, mitosis, nucleus stain, uniformity of cell, fragility and number of cells in cluster) are evaluated their possibility to be used as input data for the CBIR in order to retrieve the similar breast cases into two stages, namely malignant and benign disease.

2.2 Feature Extraction Module

The module extracts content features of submitted images. We apply cytological image Hough Transform image segmentation methods described in [13] and our work is continuing for master research [14] for Active Contour and Watershed and other radiology segmentation methods. Figure 3 shows an example of FNA segmented images. The extracted image content features are described with one set ***X={X1, …, Xn}.*** Generally ***X*** is composed of shape, colour and texture features quantified with floating-point number or vectors. So far, the features are manually extracted and described in Table 1. They are stored in features database and used to find similar cancer images with the query cancer images.

Table 1. Cytological Images Sample of Features and their Ranges

Value Range	Cytological Images X feature
Scanty, moderate or high	Cellularity
Clean , slightly dirty or dirty	Background Information
More than 21, 11-20, 6-10	Cohesiveness
31-50, 11-30, less than 10	Cell in Cluster
Present, not Present	Significant Stormal Component
Monolayered, folding, multilayered	Clump Thickness
Normal, medium, high	Nuclear membrane
Present, not Present	Bare nuclei
No, fine, prominent	Normal nuclei
Not, normal, abnormal	Mitosis
Coarse, fine	Nucleus stain
Fragile, not Fragile	Fragility

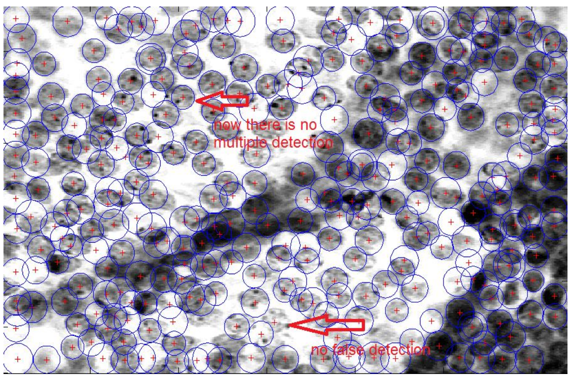

Figure 3. Result of Sobel and Circular Hough Transform

2.3 Similarity Measure

Similarity measure is one of the keys of a high performance content-based image retrieval (CBIR) system. Given a pair of images, existing similarity measures usually produce a static and constant similarity score. *Figure 4 shows sample run of query image and its retrieved images.* The Weighted Minkowski distance is used to measure (distance-) similarity between query image, **Q** and database image, I:

$$D(I,Q) = \sum_{i=1}^{M} w_i * | f_{iI} - f_{iQ} |$$

Where,

FiI and FiQ are the ith features in the query image and the database images respectively, Wi is the feature weight importance [11].

We have also tested this prototype for mobile emulator on visual studio of. Figure 5 shows sample runs of the developed prototype.

3 Experimental Results

In this section, we discuss the experimental results of our hybrid algorithm for cancer images retrieval. Our experiments are done using MATLAB integrated with Visual Studio. These are done in the framework of our research project named Mobile Cancer Management Services [MCMS]. Table 2, shows our algorithm performance. As shown, high retrieval performance is achieved for each set of cases. This high accuracy is due to the following main factors, which we fixed in our experiments:- *Usage of breast cancer wisconin dataset that is well formatted and Use of Nearest-Neighbour retrieval algorithm [11].*

Table 2: Algorithm performance

Accuracy Rate	***No. Of Test Cases***	***Average Retrieval Accuracy***
1. Benign	200	**92 %**
2.Malignant	**299**	**93 %**

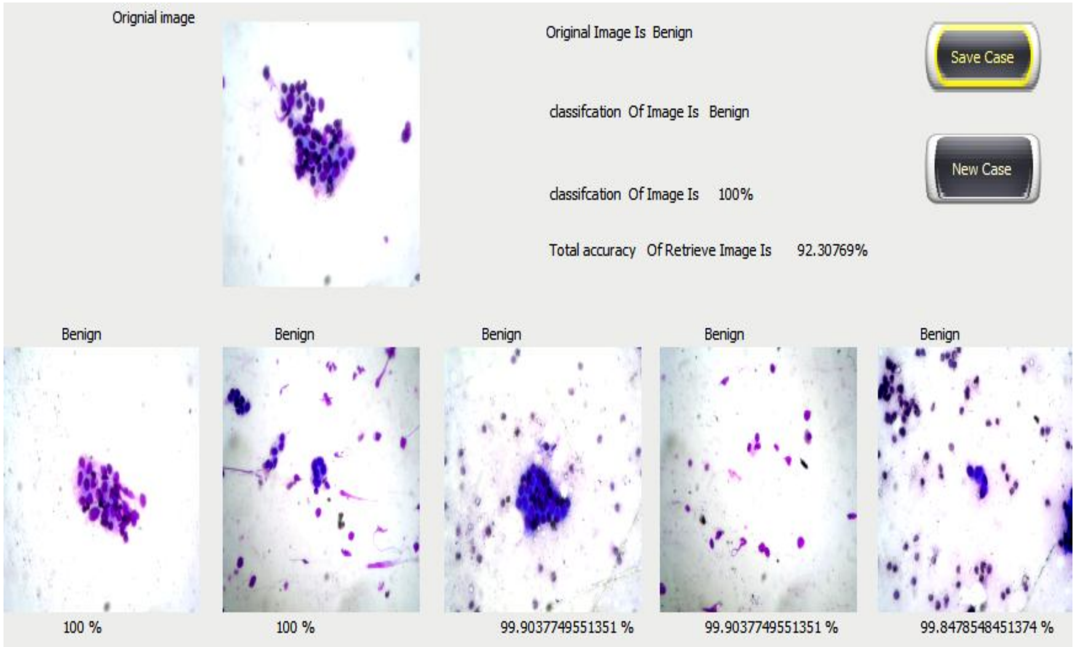

Figure 4. *Sample Run of Query Image* **Q** *and its most Similar* **K** *images*

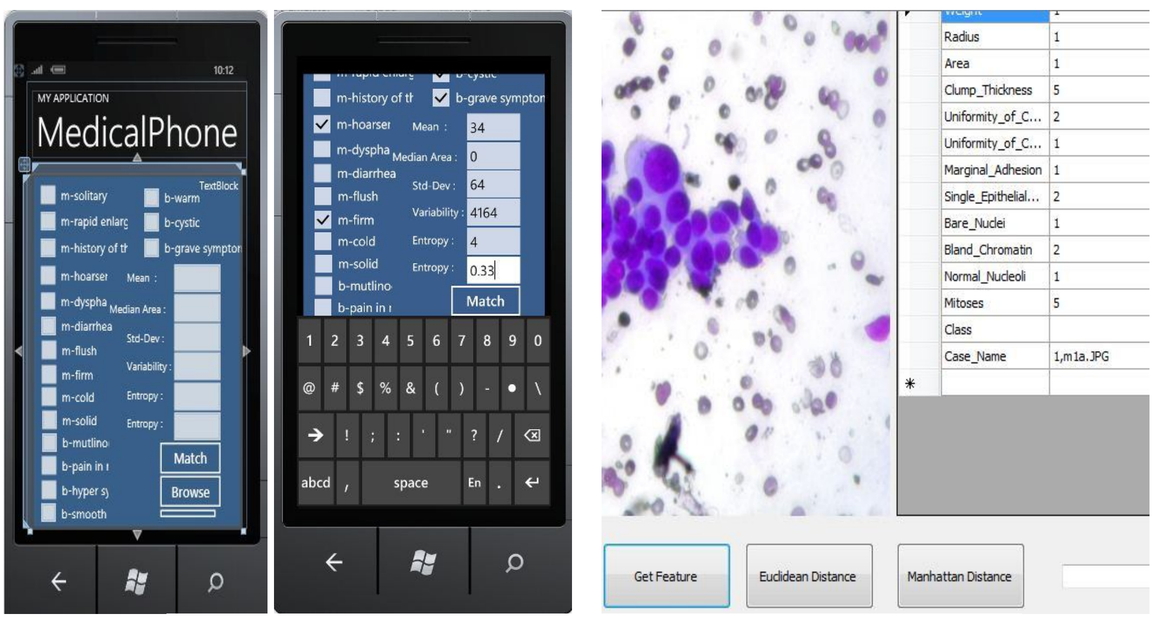

Figure 5. Sample Runs CBR and CBIR module on mobile emulator

4 Conclusion and Future Work

Our research project want to offer an interdisciplinary and international platform, namely our Mobile Cancer Management Service (MCMS) for data exchange and integration of different tools for mobile telemedicine, teleradiology, data analysis, image analysis by application of a flexible mobile platform architecture, international standards, guidelines and tools of artificial intelligence. This platform can be used as a telemedicine hospital support system as well as an interhospital support system for larger hospital associations due to the flexible system model.

In this paper, we proposed the second phase of our project, where the first phase implements a cancer expert system using case-based reasoning methodology [11]. In this second phase, we are extending the CBR retrieval module using CBIR for breast cancer FNA medical images. This is to support our mobile teleradiology component. It shows good retrieval accuracy that in average reaches 90 %. We also developed a Mobile Engine prototype for cancer remote diagnosis. In our next step of future work, we are going to implement more algorithms for image compression and enhancement algorithms for medical imaging processing on mobile phones. This is to provide more patient care and teleconsultation.

References

1. Doi K: Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput Med Imaging Graph 31(4–5):198–211, 2007.
2. Prakash M. Nadkarni, MD and Randolph A. Miller, MD, Service-oriented Architecture in Medical Software: Promises and Perils, J Am Med Inform Assoc. 2007 Mar–Apr; 14(2): 244–246.
3. Datta R, et al: Image retrieval: ideas, influences, and trends of the new age. ACM Comput Surv 40(2), 2008.
4. Muller H, et al: A review of content-based image retrieval systems in medical applications—clinical benefits and future directions. Int J Med Informatics 73(1):1–23, 2004.
5. Comaniciu D, Meer P, Foran DJ: Image-guided decision support system for pathology. Mach Vis Appl 11(4):213–224, 1999.
6. Kwak DM, et al: Content-based ultrasound image retrieval using a coarse to fine approach. Ann N Y Acad Sci 980:212–224, 2002.
7. Lim J, Chevallet J-P: Vismed: A visual vocabulary approach for medical image indexing and retrieval. in Second Asia Information Retrieval Symposium. 2005. Jeju Island, Korea.
8. Shyu CR, et al: ASSERT: a physician-in-the-loop content-based image retrieval system for HRCT image databases. Comput Vis Image Underst 75(1/2):111–132, 1999.
9. Cauvin JM, et al: Computer-assisted diagnosis system in digestive endoscopy. IEEE Trans Inf Technol Biomed 7 (4):256–262, 2003.
10. Güld MO, et al: Content-Based Retrieval of Medical Images by Combining Global Features. Accessing Multilingual Information Repositories. in Accessing Multilingual Information Repositories. 2005: Springer LNCS 4022.
11. Abdel-Badeeh M.Salem, Bassant M .El Bagoury, 2003. ***" A Case-Based Adaptation Model For Thyroid Cancer Diagnosis Using Neural Networks",*** In Proc. of the 16th International FLAIRS conference, Florida, U.S. pp. 155-159.
12. ***Breast Cancer database*** http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/breast-cancer-wisconsin.names
13. MACIEJ HREBIEN´, PIOTR STEC´, TOMASZ NIECZKOWSKI, ANDRZEJ OBUCHOWICZ. ***Segmentation of breast cancer fine needle biopsy cytological images.*** Int. J. Appl. Math. Comput. Sci., 2008, Vol. 18, No. 2, 159–170.
14. M.V.M. Figueredo and J.S. Dias, Mobile telemedicine system for home care and patient monitoring, IEEE EMBS, International Conference of the Engineering in Medicine and Biology Society, 2004, Vol. 2, pp. 3387-3390.
15. Jones SM, Banwell PE, Shakespeare PG. Telemedicine in wound healing. Int Wound J 2004;1:225—230.
16. H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathmesh Verlekar, Suraj Shirke, "Walsh Transform over Row Mean and Column Mean using Image Fragmentation and Energy Compaction for Image Retrieval", International Journal on Computer Science and Engineering (IJCSE),Volume 2S, Issue1, January 2010.
17. H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Augmented Block Truncation Coding Techniques", ACM International Conference on Advances in Computing, Communication and Control (ICAC3-2009), pp. 384-390, 23-24 Jan 2009, Fr. Conceicao Rodrigous College of Engg., Mumbai. Is uploaded on online ACM portal.