Proceedings of the 2012 IEEE 12th International Conference on Bioinformatics & Bioengineering (BIBE), Larnaca, Cyprus, 11-13 November 2012

e-Prolipsis: A web based risk estimation platform to support and register breast cancer diagnosis in Greece

Alexandros C. Dimopoulos, John Lakoumentas, Argyro Antaraki, Antonios Frigas, Emmanouil K. Ikonomakis, Marinos Sampson, Anastasios Tagaris, Aikaterini Liakou, Emmanouil Athanasiadis, George Spyrou* and Panos Ligomenides*

Biomedical Research Foundation of the Academy of Athens (*BRFAA*)
Soranou Efesiou 4, Athens, Greece.
{gspyrou@bioacademy.gr, pligo@academyofathens.gr}

***Abstract* — Breast cancer diagnosis requires specific expertise from the Medical Doctors. Furthermore, prognosis, monitoring and early detection of malignant findings can be successfully realized through synergies between physicians, researchers and general population in concert with the health policy program of each country. Information technology and computational intelligence play a crucial role in the production of digital repositories and cancer registries as well as in the development of systems to support diagnosis. In this paper we present the concept and the architecture of an approved grant under the National Strategic Reference Framework 2007-2013 (NSRF), called e-Prolipsis. Through this project, we will design and implement a web based risk estimation platform and a Central Breast Cancer Registry (CBCR) that will co-operate to provide medical doctors and patients with services such as multiplicity in diagnostic opinions, synergy, computational risk estimation, access to statistical/epidemiological data for trend estimation. Such a system, will serve as a reference tool and will help the clinician, the radiologist, the rural doctor or trainee doctor in the final assessment on the existence or likelihood of breast cancer. The system will assist in the successful diagnosis of breast cancer, giving each patient access to a large pool of doctors and at the same time, the data stored in the CBCR will be used for statistical analysis, providing useful results for improving both the cancer detection application and for making a national policy for combating breast cancer. This system will be accessible through a Web Portal, with different access levels for patients, doctors and general public and different web services available to each user group, eliminating the geographical distances that would be imposed otherwise.**

Keywords-component; large scale, breast cancer, risk estimation platform.

I. INTRODUCTION

Breast cancer is one of the most frequently occurring cancer in women and mammography is the major examination of choice since it has been proven capable to reduce the mortality of the disease [1]. However, reading a mammogram and concluding in correct diagnosis is not a trivial medical task. An early diagnosis of the disease takes the utmost importance and significantly improves the prognosis and leads to more effective treatment [2, 3]. The findings that would lead to early diagnosis and treatment are not always clearly shown – on the contrary, sometimes they are misleading. The recognition of abnormalities such as calcifications and masses in mammograms require a well-trained radiologist. Often, several findings escape the attention of the radiologist either because of fatigue or lack of experience or as a result of human error. To this end, information technology and computational intelligence have been many times employed to (i) digitally manage the images and the patient records, (ii) detect abnormalities, (iii) classify findings and provide second opinion.

A computer aided diagnosis system (CAD) that combines the aforementioned characteristics is "Hippocrates-mst", developed by our research team in the Academy of Athens. "Hippocrates-mst" is a CAD system for breast cancer which enhances the diagnostic capability of physicians by providing documented risk assessment for mammograms showing micro-calcifications [4-7]. Based on this system, we proposed a platform that would integrate services including image management, electronic patient record, risk estimation, co-examination and statistical trend report extraction, building two main infrastructures: The Central Breast Cancer Registry (CBCR) and the web based risk estimation platform to facilitate a large scale collection of images and data as well as computer aided mammographic examination for the whole female Greek population.

The proposal has been approved and the project will be implemented under the National Strategic Reference Framework 2007-2013 (NSRF), with the short name "e-Prolipsis". Hopefully it will become an instrument of prevention and early awareness of the relevant part of the population and secondly the Central Breast Cancer Registry (CBCR) will be a strong candidate to function as a National Breast Cancer Registry. It must be noted that it is the first attempt for such a registry in Greece – the CBCR will contain a huge amount of data at national level, related to the diagnosis and possibly to prognosis. These data will be derived from heterogeneous sources and consolidated into the registry.

The integrated system of diagnosis on one hand will assist in the successful diagnosis of breast cancer in collaboration

with the doctor and on the other hand will optimize and accelerate the process of medical examination for the citizens. The integrated platform will be a key structural component of a health diagnostic network, allowing each woman's medical examination to be available to an authorized group of medical personnel. In this way, multiple medical reports will be available to the woman thus accelerating the process of diagnosis and "eliminating" the geographical distances that would be imposed without the network.

II. MATERIAL AND METHODS

The platform, as already mentioned, will be web based and different access levels and services will be provided to different user groups. In order for the doctor to gain access to the platform, a special application will be freely provided, the Intelligent Digital Medical Record (IDMR), which will allow him to keep patient medical records for breast cancer. Apart from record keeping, this application will also provide the interface for the remote diagnosis of mammograms, using "Hippocrates-mst". Furthermore, a database for the CBCR will be a major component of the system, shared by all users. Finally, a web portal will be set up for awareness and population information and also for e-learning/specialization of doctors (e.g. radiologists) on critical issues of breast cancer diagnosis. All these functional components are shown in Figure 1 and are described in details in the followings.

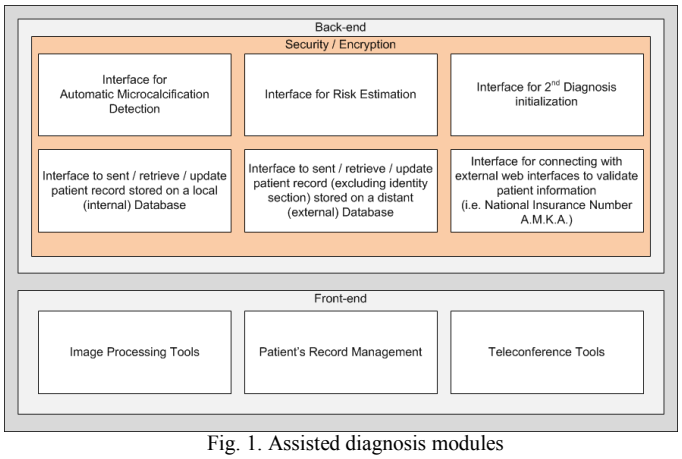


Fig. 1. Assisted diagnosis modules

A. Intelligent Digital Medical Records

The integrated diagnosis system is based on some services that will facilitate the smooth and efficient operation of the network, based on well established technology solutions and methodologies. The software implementation of the platform will be based on the principle of architecture-based Services (Service Oriented Architecture -SOA) and on Web Services. In general the model to be followed is that of the client -server as explained below:

1) Client: At the client level of the service, it can be any "consumer" in the system. However, the main consumer of the Integrated Diagnosis System is the doctor's application and the consumed services are:

- Patient's Medical Record Services
- Micro calcifications Detection Services
- Risk Estimation Services
- Assistive Diagnosis Services

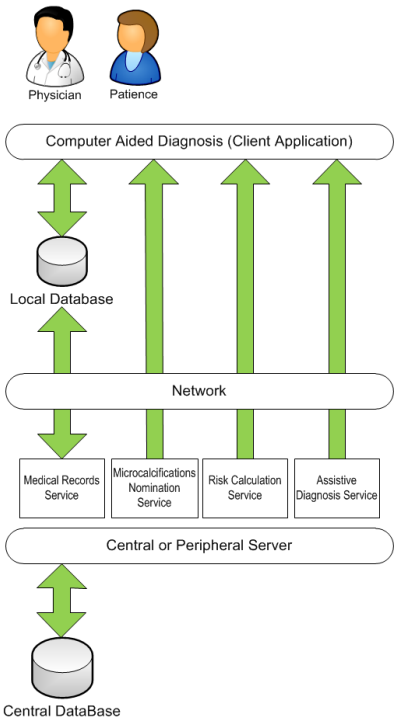
The client application will provide the doctors with various image processing tools necessary for them to select the proper region of interest (ROI) for further analysis with the interface to "Hippocrates-mst". Furthermore, an intelligent system for maintaining patient medical records for breast cancer will be implemented. The latter will keep a full medical record tailored to the needs of a mammographic examination.

The client application will be able to operate in two different modes, depending on the network connectivity. The normal functionality is when a link to the network/internet exists and communication can be established with the main server or a regional one. The second mode is when there is no such link to the network/internet. In the first case all the services of the application are available to be used by the user -doctor. In the second case, the web services of risk calculation and assistive diagnosis will not be available. However, the Patient's Medical Record Services will be available locally. In figures 2(a) and 2(a) these two modes are described by abstract representations.

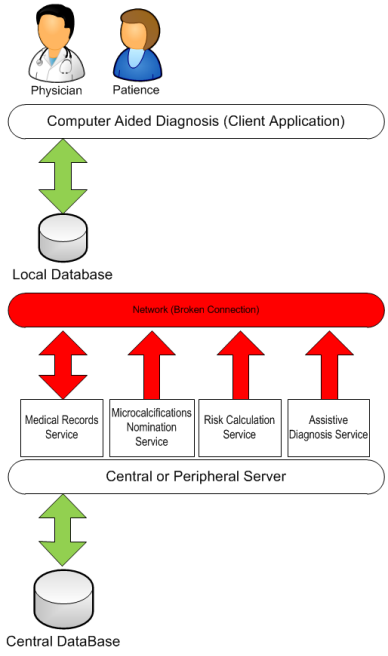
2) Server – Service Oriented Architecture: One of the major tasks of the application server is to highlight the microcalcifications and to determine the risk. These two tasks will be provided as web services, following a Service Oriented Architecture (SOA). SOA is a flexible set of design principles used during the phases of implementation and integration of computer systems. A system based on SOA provides the functionality as a sequence of services that can be used in multiple different systems.

A basic requirement of such an approach is the loose coupling of the provided services with the operating systems and other underlying technologies. SOA separates functions into services that are accessible from the internet, to allow their combination and reuse in order to build a new application. Therefore, it is important that the services and their consumers are communicating by exchanging data with a strictly defined manner/format and in compliance with predefined requirements of coordination between the available services. In the case of the proposed system, the medical records services, the microcalcifications detection, the risk identification and the assistive diagnosis constitute the basic services available to the client application.

In all the client-server transactions, security is a major issue. In order to ensure integrity and secrecy, digital signatures and encryption is used in the transactions. More specifically, the used protocols allow the usage of various types of signatures, encryption algorithms and security tokens (X.509 certificates, Kerberos tickets, UserID/Password credentials, SAML-Assertion, Custom defined token).



(a) On-line Mode



(b) Off-line Mode

Fig. 2. Abstract representation of the two operation modes.

B. *Computer Aided Diagnosis*

Our group at the Academy of Athens has designed and implemented a system of Computer Aided Diagnosis (CAD), called "Hippocrates-mst" [4-7]. This system includes a "smart" system of registration of patients, that uses and applies well known risk models, producing a complete diagnosis for breast cancer patients. "Hippocrates-mst" urges the doctor to focus on the region of interest (ROI) and to analyze it with the provided digital tools. Then, the system returns to the doctor the number and type of detected micro calcifications as well as the risk percentage attributed to that region. That is, unlike the other commercially available softwares, "Hippocrates-mst" is interactive and actively involves the doctor in the process.

The system operation can be described in the following simple steps:

1) The patient's mammogram (if not already digital) is digitized (with a proper mammogram scanner) and then imported into the patients' record. If it is a new patient, a new record is created.
2) The doctor examines the mammography images, using the digital tools provided by Hippocrates-mst, trying to identify and classify suspicious areas. These areas are marked as ROIs and are forwarded to "Hippocrates-mst" analysis tool.
3) "Hippocrates-mst" outputs its results coloring each area with a different color depending on the risk. Finally, the overall risk of breast cancer is presented.
4) The doctor can either further examine other ROIs or choose to save the results in the patients' record and terminate the process.

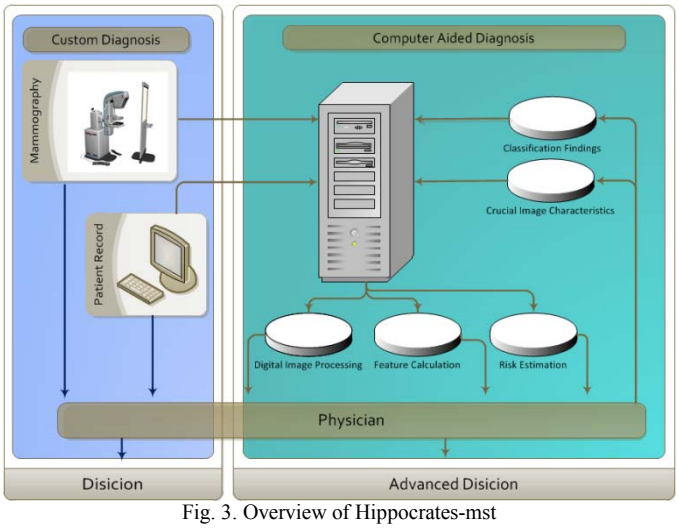


Fig. 3. Overview of Hippocrates-mst

Therefore, depending on the outcome level of risk, the patient is categorized on a scale like that of the BI-RADS [8] system, but less detailed (1 – acquittal, 2 – biopsy suggested, 3 – biopsy recommended, 4 – biopsy required). In short, the diagnostic system of "Hippocrates-mst" offers the physician a friendly work environment and helps him extend the diagnostic capability by providing overview tools on the digital/digitized images. The physician role in the diagnosis is active, allowing the step-by-step focus on dictated suspicious areas.

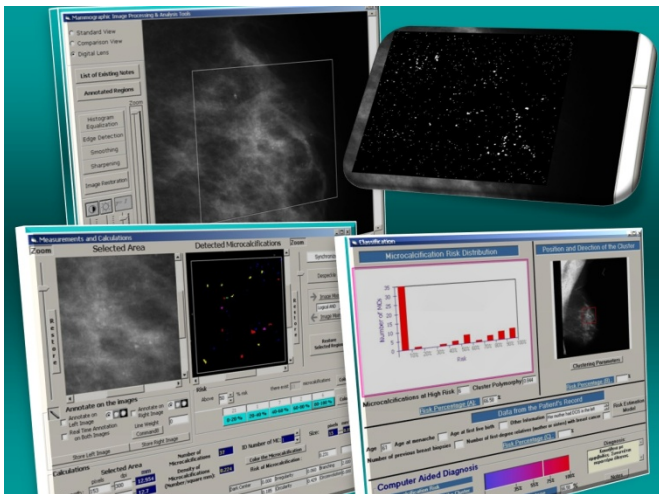


Fig. 4. Snapshots from Hippocrates-mst

The e-Prolipsis platform is going to be based on "Hippocrates-mst" and the abovementioned basic principles are presented in figure 3 where as snapshots of "Hippocrates-mst" are shown in figure 4. The computer aided diagnosis system will take as input mammographic images, which will be processed digitally, and patients' historical data along with values of various risk factors, which will be integrated in the risk assessment. During this process, the direct supervision and intervention of doctors will be of major importance. The doctor will be invited to combine the computational findings with his/her human knowledge and experience to conclude to an optimum diagnosis.

C. CBCR Databases

The main objective of the CBCR is the collection of data and findings regarding breast cancer nationwide, in a database accessible by collaborating physicians, scholars and by ordinary users interested in the situation of breast cancer in the country. The data stored in the local databases of the participants will be gathered in CBCR, after securing both the patients' consent and the removal of any information leading to identifying the patient associated with them. The data stored in the CBCR will be used for statistical analysis to determine the risk parameters of mammographic findings and to record similar events based on various criteria of the patients' record. The results of the above process will be properly formatted in a readable format and will be available through a secured web site to properly authorized users. The results that emerge from the statistical analysis, will be used to improve both cancer detecting applications and reporting of the medical staff to take into account in future reports various factors that will arise. These statistical results, in the form of epidemiological studies will be published in the Portal of the e–Prolipsis and some of them will be sent to all registered doctors in the network. The proposed central database will:

- Collect, maintain and organize the data in the Registry.
- Provide easy and reliable access to parts of the system that make use of data.
- Enable the extraction of useful conclusions concerning the database contents.

Apart from the central database, additional local databases will be developed for the doctor client-application that will synchronize remotely with the central database in a transparent mode for the doctors.

The CBCR database will be installed in the Biomedical Research Foundation of the Academy of Athens (BRFAA) premises, at dedicated DB Servers on provision for high availability. The DB system will provide graphical interface for central control and management for database administrators and will support technologies for maximizing the utilization efficiency as much as possible. The maximum number of doctors expected to register in the system is 3,000, of which 500 it will be feasible to be concurrently online. The overall system will be able to maintain up to 2,500,000 records – the population eligible to take mammography (women aged 40-79).

To ensure the data credibility and integrity the database management system (DBMS) will operate in an environment in which data are encrypted and their integrity is protected when they are transmitted from the server, using an appropriate standard such as the SSL. The stored data will be encrypted using encryption algorithms such as DES (56-bit keys), 2-key 3DES (112-bit keys), 3-key 3DES (168-bit keys) or equivalent.

Moreover, simultaneously access of multiple users on the same data will be supported by ensuring consistency in data reading, integrity and accuracy, without any delay to the user.

Finally, in order to ensure data security, the DBMS will have a complete mechanism for monitoring and recording the use of database users (auditing), both for successful and unsuccessful operations.

III. USAGE SCENARIO

In order to enlighten system utilization, a typical usage scenario is presented. Initially the doctor–user connects to the e-Prolipsis Portal and registers the required personal data and credentials for the system. The successful registration will allow the user to download the application for the Intelligent Digital Medical Records (IDMR) and also provide him with access to the CBCR, to upload the existing medical records.

When a patient visits a registered doctor for the first time a new record will be created locally on the doctor's IDMR application – next time she visits the same doctor all new medical information will be appended to the existing record. The doctor scans the mammography and stores the digital image locally. Using the IDMR image processing tools, the original image is processed and after a ROI is defined, the image is sent through the internet to the "Hippocrates-mst" for analysis. The results are sent back to the doctor and simultaneously the medical

record is also stored into the CBCR. From this point, the doctor can proceed with his/her diagnosis which he/she stores on the patient's medical record. In case the doctor wants a second opinion on the mammography, he can simply activate the option for second diagnosis by another doctor, registered to the system. In this way, the second doctor may examine through the CBCR the patient record and the "Hippocrates-mst" output and he is asked to store his diagnosis in the patient's medical record. Once a patient has already been registered to the CBCR by a doctor, she can also trigger the second opinion option by herself, from a doctor of her choice. The results from the second doctor can either be received by mail at her address or at her profile in the e-Prolipsis platform. Furthermore, based on the medical records in the CBCR, the patient can choose to receive alerts via sms or e-mail on topics such as notification for repeat examinations, the estimated time for medical exams results, news and updates related to breast examination.

To give the opportunity for each patient to seek diagnosis from a doctor who participates in the network regardless of geographic constraints, it will be possible to update or even register a patient to the CBCR through dedicated service centers for patients. In such a center the patient, with the help of trained personnel, will be able to digitalize her mammography and create a medical record in the NBCR. The present grant will cover the expenses of 11 such centers distributed over Greece.

Last but not least, we should mention that through the e-Prolipsis portal digital information and advice will be provided to citizens on issues related to the prevention of breast cancer. Especially for register users, more specific information will be provided. Finally, e-learning services will be available for doctors in specialized subjects of breast cancer diagnosis.

IV. ACKNOWLEDGMENTS

This work is funded by the operational program "Digital Aid" under the National Strategic Reference Framework 2007-2013 (NSRF), with title: "Development of a DataBase for the National Breast Cancer Registry and implementation of a web platform for tele-diagnosis and preventive monitoring of female population aiming to the early diagnosis of breast cancer, e-Prolipsis", with code number MIS 296160.

REFERENCES

[1] American Cancer Society. Global cancer facts and figures, 2007.

[2] Lehman CD Elmore JG, Armstrong K and Fletcher SW. Screening for breast cancer. Journal of the American Medical Association, 293(10):1245–1256, 2005.

[3] McCallum M, Baker Cand Gillespie K, Kohen B, Stewart H, Leonard R, Cameron D, Leake R, Paxton J, Robertson A, Purdie C, Gould A, and Steel M. A prognostic index for operable, node-negative breast cancer. British Journal of Cancer, 90(10):1933–1941, 2004.

[4] Spyrou G, Kapsimalakou S, Frigas A, Koufopoulos K, Vassilaros S, and Ligomenides PA. "Hippocrates-mst": a prototype for computer-aided microcalcification analysis and risk assessment for breast cancer. Medical and Biological Engineering and Computing, 44(11):1007–1015, 2006.

[5] Frigas A, Kapsimalakou S, Spyrou GM, Koufopoulos K, Vassilaros S, Chatzimichael A, Mantas J, and Ligomenides PA. Evaluation of a breast cancer computer aided diagnosis system. Medical Informatics in a United and Healthy Europe, 124:631–636, 2006.

[6] Frigas A, Spyrou GM, Antaraki A, Patiraki E, Koufopoulos K, Mantas J, and Ligomenides PA. Design of a 'smart' patient record system for mammography patients. Medical Informatics in a United and Healthy Europe, 150:130–134, 2009.

[7] Argyro Antaraki Ioannis Andreadis Dimitra Koulocheri Flora Zagouri Afroditi Nonni George M. Filippakis Konstantina S. Nikita Panos A. Ligomenides Georgia Giannakopoulou, George M. Spyrou and George C. Zografos. Downgrading birads 3 to birads 2 category using a computer-aided microcalcification analysis and risk assessment system for early breast cancer. Computers in biology and medicine, 40(11):853–859, November 2010.

[8] Liberman L and Menell JH. Breast imaging reporting and data system (bi-rads). Radiologic clinics of North America, 40(3):409–430, May 2002.