Cloud based Decision Support System for Diagnosis of Breast Cancer using Digital Mammograms

S. Aruna
Research Scholar
Dept of Computer Applications
Dr MGR Educational and
Research Institute University

L. V. Nandakishore
Asst Professor
Dept of Mathematics
Dr MGR Educational and
Research Institute University

S. P. Rajagopalan, PhD.
Prof emeritus
Dept of computer science
Dr MGR Educational and
Research Institute University

ABSTRACT
In this paper, we propose a cloud based decision support system for screening breast cancer using digital mammograms. The proposed system is deployed in a private cloud as software / infrastructure as a service. The combination of image enhancement techniques, feature extraction techniques, feature selection techniques, ensemble neural networks for classification, results verification process and deployment in the private cloud are added advantages for effective performance of the system.

General Terms
Image processing, Expert systems.

Keywords
Breast cancer, Digital Mammograms, Neural networks, Ada boost, Feature extraction, Feature selection, Cloud computing.

1. INTRODUCTION
Cancer detection has become a significant area of research in pattern recognition community. Breast cancer is one of the leading causes of cancer mortality among women in the United States [1]. It causes significant worry among the women and their physicians. Although breast cancer is a potentially fatal condition, early diagnosis of disease can lead to successful treatment [2]. Early diagnosis needs a precise and reliable diagnostic procedure that allows physicians to distinguish between benign breast tumors and malignant ones [3].

Mammography using low energy X-ray of human breast called mammograms are the first step in screening and detecting breast tumors. Digital mammography is a specialized form of mammography that uses digital receptors and computers instead of x-ray film to help examine breast tissue for breast cancer [4]. Digital mammograms are better at detecting early stage breast cancer [5]. However the sensitivity of mammography varies with an increase in breast density. Independent double reading by two radiologists has been shown to improve the sensitivity, but it also increased the cost of the screening process [6]. Sometimes the radiologists for second opinion may be unavailable. Hence there is a need for automated systems for decision making.

Clinical decision support systems (CDSS) assist physicians in the detection of diseases thus reducing errors in diagnosis. There are two types of CDSS, namely those using knowledge base and inference engine and those using machine learning algorithms. Machine learning algorithms based systems is fast and effective for a single disease. They can give a second opinion to the physicians to aid in selecting the treatment strategies without human intervention. Some of the limitations of these image processing systems are they are standalone systems, need lots of storage and processing powers and increased cost of maintenance. These boundaries can be prevailed by providing a CDSS in the cloud.

In this paper we propose a breast cancer decision support system (BCDSS) to diagnose breast cancer from digital mammograms using the cloud. The system is projected to guide radiologists and physician in the breast cancer decision-making process. The rest of the paper is organized as follows. Section 2 explains the proposed System architecture and the conclusions and future work are provided in the section 3.

2. CLOUD BASED BCDSS SYSTEM
The cloud offers hardware and software services as virtualization of resources on the internet managed by third parties. These services include advanced software applications and high end networks and servers. These services are offered to the end user without the necessity of knowledge of the systems utilized by them.

CDSS for screening breast cancer using digital mammograms need high performance image processing systems, expertise persons for operation and maintenance. They are also stand alone and expensive systems. When CDSS are offered in the cloud end users are benefited with reduced cost, increased storage, flexibility, portability, scalability, maintenance etc. Additionally as the data is stored in the cloud the data is not lost during the shifting of hospitals, hardware errors, earthquake etc.

2.1 BCDSS
Breast cancer decision support system BCDSS is deployed in a private cloud as software / infrastructure as a service for a particular platform i.e. health care in the breast cancer domain. Private clouds offer a higher level of security and regulatory compliance than most public cloud implementations. This enables high security and control over the data. The end users have no worry about the updating, upgrading or maintenance of the system. BCDSS is an image processing system processes digital mammograms. Figure 1 shows the architecture of the BCDSS system.

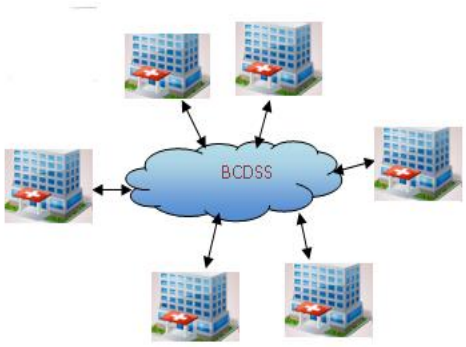

Fig 1 ; BCDSS architecture

The hospitals/testing centers registered in the cloud can store the digital mammography images in the virtual storage area. These images can be processed either in individual or batch mode. If batch mode is selected the loop will continue processing until no mammograms remain for processing. Figure 2 shows the stages in BCDSS

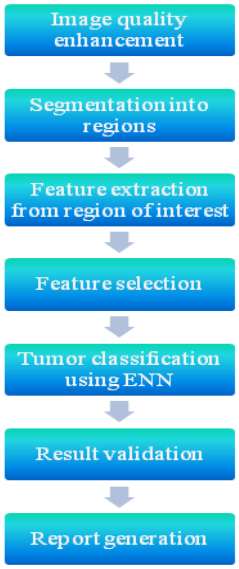

Fig 2; Stages in BCDSS

The sequence of steps in BCDSS is as follows:

Step 1: Image preprocessing techniques like smoothing, sharping, pectoral muscles removal are used to improve the quality of the mammograms.

Step 2: The resultant image is segmented into different regions.

Step 3: Features are extracted from the region of interest.

Step 4: From the extracted features using feature selection techniques a subset of the best features are selected for further processing.

Step 5: Classification of selected features using Ada boost with neural networks as base classifier.

Step 6: The classification results are validated using performance metrics

Step 7: Reports are generated for the processed mammograms.

The generated reports are stored in the data center of the server which can be viewed/printed by the hospitals. As each hospitals' data and reports are stored in separate files the hospitals who own the data only can access them after authentication. Therefore data and reports will be secured from unauthorized access. High accuracy of the system will be maintained by cross validation of the results. Only the best features from the region of interest are used for classification to enable faster performance. Artificial neural networks (ANN) are used as the base classifier. The benefits of neural networks namely easy implementation, ability to learn and generalize the patterns similar to humans, adaptability in learning, distribution of knowledge in the entire network, effective learning even features in high dimensional space and ability to predict correctly even with noisy irrelevant features have made us to choose ANN as a base classifier for the system. In addition, the Ada boost technique of ensemble approach is used to improve the predictive accuracy of the ANN. Ada boost algorithm, first introduced by Freund & Schapire improves the performance of any given classifier.

2.2 Advantages of BCDSS

- The system can be used either in single or batch processing mode
- The system is deployed in a private cloud computing environment as Software / infrastructure as a service.
- Due to the deployment in the private cloud high security for the patients' data.
- Only authenticated users can have access to the bcdss system.
- Image enhancement techniques, feature extraction and selection techniques, validation metrics together with ensemble classifiers yield a high accuracy, performance and cost effective system.
- As the system is available in the cloud, BCDSS ensure portability, scalability and flexibility.
- No special maintenance / need for expert personnel for the system
- The system reduces the workload for physicians and radiologists.

3. CONCLUSION

Breast malignancy is one of the most common cancer among women. Early detection of the disease can improve the long time survival rate and selecting appropriate treatment methodologies for the disease. Digital mammograms are found to be effective for screening and diagnosing breast cancer. In this paper we propose a cloud based Breast cancer decision support system (BCDSS) for diagnosing breast cancer using digital mammograms. The system is deployed in a private cloud computing environment and implemented as software / infrastructure as a service to ensure data security and to support both novice and expert users. The system uses image enhancement techniques to improve the quality of mammograms. Then mammograms are segmented into regions and features are extracted from the region of interest.

From the extracted features a feature subset of the best features is selected using feature selection techniques. The selected features are used for classification using ensemble neural networks. The results are validated and reports are generated. The system is intended to assist physicians and radiologists without human intervention thus saving expert human resources. Future work will be concentrated in evaluation of the system which will help in further improving of the system.

4. REFERENCES

[1] Breast Cancer What Are the Key Statistics for Breast Cancer?. American Cancer Society Cancer Resource Information. http://www.cancer.org

[2] Harirchi, et al., ¯Breast cancer in Iran: a review of 903 case records,. Public Health, 2000. 114(2): p. 143-145.

[3] Subashini. T, Ramalingam. V, Palanivel. S, 2009, ¯Breast mass classification based on cytological patterns using RBFNN and SVM,. Expert Systems with Applications, 36(3): p. 5284-5290.

[4] Sariego J, 2010. "Breast cancer in the young patient". The American surgeon, **76** (12), pp 1397–1401

[5] Kekre HB, Sarode Tanuja K and Gharge Saylee M, 2009, "Tumor Detection in Mammography Images using Vector Quantization Technique", International Journal of Intelligent Information Technology Application, 2(5):237-242.

[6] Baines CJ, McFarlane DV, Miller AB, 1990, "The role of the reference radiologist: Estimates of interobserver agreement and potential delay in cancer detection in the national screening study". Invest Radiol, 25: 971-076.