

UGC MAJOR RESEARCH PROJECT

FINAL REPORT



Content Based Image Retrieval Coupled Computer Aided Diagnosis of Breast Cancer in Digital Mammograms

Principal Investigator : **Dr. T. S. Subashini,**
Associate Professor,
Department of Computer Science and Engineering,
Annamalai University,
Annamalainagar,
Tamil Nadu - 608 002.

CO-Investigator : **Dr. M. Kalaiselvi Geetha,**
Associate Professor,
Department of Computer Science and Engineering,
Annamalai University,
Annamalainagar,
Tamil Nadu - 608 002.

UNIVERSITY GRANTS COMMISSION

BAHADUR SHAH ZAFAR MARG

NEW DELHI – 110 002

**PROFORMA FOR FINAL REPORT OF THE WORK DONE ON THE
PROJECT**

1. TITLE OF THE PROJECT : Content Based Image Retrieval Coupled
Computer Aided Diagnosis of Breast Cancer in
Digital Mammograms.

2. NAME AND ADDRESS OF THE PRINCIPAL INVESTIGATOR :

Dr. T. S. SUBASHINI
Associate Professor,
Department of Computer Science & Engineering,
Faculty of Engineering & Technology,
Annamalai University.

3. NAME AND ADDRESS OF THE INSTITUTION :

Department of Computer Science & Engineering
Faculty of Engineering & Technology
Annamalai University

4. UGC APPROVAL LETTER NO. AND DATE :

F. No. 41-635//2012 (SR) dated 16th July 2012

5. DATE OF IMPLEMENTATION : 01.07.2012

6. TENURE OF THE PROJECT : From 1.7.2012 to 30.06.2015

7. TOTAL GRANT ALLOCATED : Rs. 11, 31, 800

TOTAL GRANT REALLOCATED : **Rs. 11, 02, 381**

8. TOTAL GRANT RECEIVED : Rs.10, 33, 023

9. FINAL EXPENDITURE : Rs. 10, 32, 857 + 69, 358 = **Rs. 11, 02, 215/-**

[**Rs. 10, 32, 857/-** : Expenditure from the Grant released, **Rs. 69, 358/-** : Expenditure incurred against the third installment amount, to be released by UGC].

10. TITLE OF THE PROJECT : Content Based Image Retrieval Coupled
Computer Aided Diagnosis of Breast
Cancer in Digital Mammograms

11. OBJECTIVES OF THE PROJECT

- To create a Content Based Mammogram Retrieval framework based on breast tissue.
- To create a Content Based Mammogram Retrieval framework based on pathology.

12. WHETHER OBJECTIVES WERE ACHIEVED

YES

The Content based retrieval system would aid radiologist to make accurate diagnosis by displaying mammograms having similar pathologies and past diagnosis. Further it can be useful as a teaching aid for medical and PG students of Radiology.

13. ACHIEVEMENTS FROM THE PROJECT

The greatest achievement from this project is the successful development of the Content based Mammogram Retrieval (CBMR) system for aiding the radiologist in diagnosis and decision making. It can be used to teach and train medical students and PG students of radiology by displaying similar images of various patients to learn and interpret specific pathologies when resources are limited. The final output of the project was exhibited to radiologists of the Medical Imaging Department of Raja Muthiah Medical College and Hospital housed in our Institution and the feedback from them was positive and encouraging.

14. SUMMARY OF THE FINDINGS

In this project work, methods were proposed for content based mammogram retrieval based on breast tissue and breast pathology using image processing techniques and pattern classification methods. The development of the content based mammogram retrieval system comprises of the following major tasks namely

- i. Preprocessing
- ii. Segmentation and Feature Extraction
- iii. Modeling for database creation
- iv. Retrieval based on content

In the preprocessing step, artifacts present in the mammogram and pectoral muscles were eliminated. A new algorithm for eliminating pectoral muscle in digital mammograms was proposed. These procedures served as a preprocessing step for subsequent analysis of mammograms for detection and diagnosis of breast cancer. It significantly reduced the number of pixels that need to be used in the extraction of features for classification. Results and examples of the outstanding performance of these algorithms were presented.

Three different methods for ROI segmentation is proposed and the best method which gives the better classification performance is utilized in the proposed CBIR framework based on breast tissue density. Statistical and Haralick features are derived from the region of interest to classify the breast density using SVM into any of the three classes namely dense, glandular and fatty. A separate database is created which consists of images and its corresponding features for each of the tissue class. This helps to restrict the search for retrieval to a limited number of images relevant to the class of the query image. k-NN with three different distance measures are used for retrieval and the performance of each of the distance measures were reported and compared. The results indicate that Haralick features modeled using SVM and retrieval done using k-NN with Chebychev distance performed well when compared to others.

Also an pathology based CBMR is proposed in this work. Wavelet features were modeled with radial basis function neural network and support vector machine to classify normal and abnormal mammograms. The best performance was achieved with features extracted from wavelet decomposition level 2, because mass and microcalcifications are represented as high frequency information which were obtained in the highest wavelet decomposition levels.

Next the abnormal mammograms alone are considered for pathology/mass detection. Minimum cross entropy segmentation approach for detecting the

suspicious regions from the whole breast region is carried out. After segmenting the suspicious regions, Local binary pattern, geometrical and Zernike moment features are derived and model based techniques SVM, Adaboost and LDA is used to discriminate the mass and false positives from the suspicious regions.

Further, CBIR framework for mammograms based on the type of severity of the mass is proposed. Geometrical features and Zernike moment features using SVM, Adaboost and SRC classifiers are used for classifying the benign and malignant mass. SVM with geometrical features results outperforms and it can be effectively used in the diagnosis of mass type severity, so these classified images and its features are used for CBIR framework which retrieves the most relevant images closest to query image based on type of severity. k-NN with three different distance measures are used for retrieval and the performance of each of the distance measures were reported and compared. The results indicate that Geometrical features modeled using SVM and retrieval done using k-NN with Chebychev distance performed well when compared to others.

15. CONTRIBUTION TO THE SOCIETY

The societal benefits of this project is high and few of them are

- i. It can be used as a second reader by the Radiologist to make accurate diagnosis and decision making thereby reducing false positives and false negatives.
- ii. It can be used as teaching aid to train medical students in interpreting mammogram images.

16. WHETHER ANY Ph.D. ENROLLED/PRODUCED OUT OF THE PROJECT

Ph.D Enrolled : YES

Candidate Name : K. Vaidehi

Enrolled on : November 2011.

Status : Synopsis submitted.

17. NO. OF PUBLICATIONS OUT OF THE PROJECT

S. No.	Paper Title	Name of the Journal/ Conference	Status
1	Automatic Characterization of Benign and Malignant masses in Mammography	Procedia Computer Science (Elsevier)	Published during April 2015
2	Breast Tissue Characterization using combined k-NN classifier	Indian Journal of Science and Technology	Published during Jan. 2015
3	An Intelligent Content Based Image Retrieval System For Mammogram Image Analysis	Journal of Engineering Science and Technology	Accepted on Feb. 2015 & will be published during Nov 2015
4	Automatic Classification of CC view and MLO view in Digital Mammograms	Lecture Notes in Electrical Engineering (Springer)	Published during Nov. 2014
5	Automatic Detection of Suspicious Region and Removal of Pectoral Muscle using Minimum Cross Entropy Thresholding	2 nd International Conference on Emerging Research in Computing, Information, Communication and Applications (Elsevier proceedings)	Published during Aug. 2014
6	Content Based Benign and Malignant Mass Mammograms Retrieval	International journal of Applied Engineering Research	Published during Dec. 2014
7	Automated Characterization of Mammographic Density for Early Detection of Breast Cancer Risk	International Journal of Simulation, Systems, Science and Technology	Published during Feb. 2014
8	Automatic Identification and Elimination of Pectoral Muscle in Digital Mammograms	International Journal of Computer Applications	Published during Aug. 2013
9	A global approach for detecting mass in Digital Mammograms	International Journal of Advancements in Research and Technology	Published during Nov. 2012