INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

# SURVEY ON COMPUTER AIDED DIAGNOSIS METHODS FOR BREAST CANCER DETECTION AND CLASSIFICATION

**Anusha D.S.**

PG Full Time Student, CSE Department, Karunya University, Coimbatore

anudeepa125gmail.com

**Abstract**

Mammography is used to detect breast cancer in women. Breast cancer may be undetected by the radiologist because of mammographical antiquity. This study provides computer aided diagnosis (CAD) methods for automatic detection of breast cancer. CAD provides easy detection of breast cancer by radiologist. The CAD has pre-processing and classification steps for identification of breast cancer.

**Keywords**: Mammography, image segmentation, feature extraction, cancer classification

## 1. Introduction

Mammography is used to examine human breast and is used as a diagnostic tool. The advantage of mammography is early detection of masses. Mammography remains difficult in some places to detect suspicious region of interest. Region of interest is segmented to analyze abnormalities. The segmented region of interest is then classified to mass or non mass. CAD helps to detect true positive masses and remove false positives. The steps in breast cancer detection using CAD are

1. Pre-processing

2. Image segmentation

3. Feature extraction

4. Mass/non-mass classification

In pre-processing step, the image is enhanced. It adjusts the contrast and brightness for the mammogram image. Filter function is used to enhance the image. CAD provides techniques that will automatically detect mass in mammogram. In segmentation, mammographical image is partitioned into regions. In feature extraction, input data is transformed into set of features. In classification, the mammogram image is classified into mass or non mass.

## 2. Methodology

The diagnosis begins with pre-processing. Then segmentation is done in order to partition the regions. Then features are extracted from the regions. Then, the classification of mass will be done.

### 2.1 Segmentation

Segmentation is the process of partitioning an image into non overlapping regions. In segmentation, the detailed description of the image will be given as output. The segmentation methods are described below.

#### 2.1.1 Watershed segmentation

Watershed segmentation [9] algorithm segments regions into Catchment basins. The term watershed refers to a ridge that divides areas drained by different river systems. A catchment basin is the geographical area draining into a river or reservoir. In watershed algorithm, first read the input greyscale image. Then use the gradient magnitude function as segmentation function and mark the foreground objects. Compute the background markers and compute the watershed function and visualize the result.

#### 2.1.2 Level-Set Segmentation

Level-set [2] methods start with an initial region and evolve this region while minimizing region energy. Region energy can be calculated by using level-set equation. The algorithm stops when it reaches the required threshold value.

#### 2.1.3 Local Seed Region Growing Segmentation

In local seed region growing [1], seed selection is based on local and global conditions. First apply seed criterion to all the pixels of the image. Select the pixels that satisfy the condition as seed. Start growing the region from the seeds.

### 2.2 Feature Extraction

Feature extraction is used to transform input data into set of features. Feature extraction can be based on size and geometric shape of the selected image. Some of the feature extraction techniques are described below.

#### 2.2.1 Discrete Wavelet Transform

In discrete wavelet transform [3], wavelets can have discrete values. It captures location and frequency information. Haar wavelet coefficients are used for image representation.

#### 2.2.2 Over complete Wavelet Transform

Over complete Wavelet Transform [3] gives higher number of wavelet coefficients. Redundant encoding of image information is performed by over complete wavelet transform. It uses Haar wavelet coefficient for image representation.

### 2.2.3 Spherical Wavelet Transform

Spherical wavelet transform [1] is used in detection of features in spherical shaped images. It provides multi-resolution analysis. Features extracted can be relate to size, geometric shape and boundary contour from spherical wavelet transform coefficients. The features can be calculated to provide feature metrics.

### 2.3 Mass Classification

The extracted features can be then classified into mass or non mass. The methods used for classification are described below.

### 2.3.1 Artificial Neural Networks

It is a classifier used to classify mass. It is a parametric method. Artificial Neural Networks (ANN) [4] consists of three parameters, the interconnection pattern between different neuron layers, the learning process for updating weights and the activation function that converts input into output. ANN is capable of machine learning and pattern recognition.

### 2.3.2 K-Nearest Neighbour

K-Nearest Neighbour (KNN) [4] is a non parametric method used for classification. It is an instance-based classifier. It consists of training and validation data set. Training data set consists of neighbouring proximity of input. Validation data set consists of output class.

### 2.3.3 Cognitive Resonance

In cognitive resonance [5], a context-free grammar is developed, to identify the mass type. It is a simple and low cost classification function. It generates discrete linguistic output.

### 2.3.4 Soft Cluster Neural Network

The soft cluster [6] is used in neural network to improve the generalization ability of neural network. The features were clustered into soft clusters. The weak clusters can be removed and the soft cluster neural network is trained by calculating the weights of the network. Then it can be tested. Soft clusters avoid iterative processes.

### 2.3.5 Bayesian Networks

Bayesian network (BN) [7] consists of qualitative part and quantitative part. Quantitative part gives the structural model and qualitative model gives the set of local probability distributions. Bayesian network is a direct acyclic graph. The BN classifiers are Naïve Bayes classifier, Bayes-N classifier, MP-Bayes classifier etc.

### 2.3.6 Support vector based fuzzy neural network

The support-vector-based fuzzy neural network (SVFNN) [8] takes advantage of classification of support vector machine (SVM) and the efficient reasoning of FNN in handling uncertainty information. It consists of three phases, learning phase 1, learning phase 2 and learning phase 3. Learning phase 1 consists of four layers, layer 1 consists of inputs, layer 2 consists of functions, layer 3 consists of the logic rules and layer 4 is the output layer. In learning

phase 2, input is partitioned and weights are optimized. In learning phase 3, unnecessary fuzzy rules are eliminated. Leave-one-out cross validation method is used for validating the correctness of test output.

### 2.3.7 Cross-Validation

Cross-Validation [3] is a technique used to train and test data. It is applicable when the dataset contains limited number of instances. It divides the dataset into n folds. Each fold contains equal number of instances. One fold will be kept for testing and all other folds are used for training set. Each fold will be tested once.

### 2.3.8 Support Vector Machine

Support Vector Machine (SVM) [1] is a method used to estimate data classification function. SVM constructs a hyper plane as the margin of separation between positive and negative results. The feature matrices are given as input to SVM. Each instance in the training set consists of several attributes and one output class. SVM uses kernel function. Kernel function is used to analyze patterns.

## 3. CONCLUSION

Computer Aided Diagnosis (CAD) provides techniques for automatic detection of breast cancer. It gives benign/malignant classification. It reduces the number of unnecessary biopsies.

## REFERENCES

[1] P. Görgel et al., Mammographical mass detection and classification using Local Seed Region Growing–Spherical Wavelet Transform (LSRG–SWT) hybrid scheme, Computers in Biology and Medicine 43 (2013) 765–774

[2] T. Berber, A. Alpkocak, P. Balci, O. Dicle, Breast mass contour segmentation algorithm

in digital mammograms, computer methods and programs in biomedicine 110 (2013) 150–159

[3] E. Angelini, R. Campanini, et al., Testing the performance of image representations for mass classification in digital mammograms, Int.J.Mod.Phys.C17 (2006)113–131.

[4] P. Gorgel, A. Sertbas, O.N. Ucan, A comparative study of breast mass classification based on spherical wavelet transform using ANN and KNN classifiers, Int.J.Electron.Mech.Mechatronics2(2011)79–85.

[5] A. Tahmasbi, F. Saki, S.M. Seyedzadeh, S.B. Shokouhi, Classification of breast masses based on cognitive resonance, in: Proceedings of the IEEE, 3rd International Conference on Signal Acquisition and Processing (ICSAP'2011), Singapore, 2011, pp. V1-97–V1-101.

[6] B. Verma, P. McLeod, A. Klevansky, A novel soft cluster neural network for the classification of suspicious areas in digital mammograms, Pattern Recognition 42 (9) (2009) 1845–1852.

[7] N. Ramirez, H. Acosta-Mesa, H. Carillo-Calvert, L. Nava-Fernandez, R. Barrientos- Martinez, Diagnosis of breast cancer using Bayesian networks: a case study, Computers in Biology and Medicine 37 (2007) 1553–1564.

[8] F. Moayedi , Z. Azimifar, R. Boostani,S. Katebi, Contourlet-based mammography mass classification using the SVM family, Computers in Biology and Medicine 40 (2010) 373–383.

[9] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE Transactions on Pattern Analysis and Machine Intelligence

13 (6) (2002) 583–598.