ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 13, March 2016

PREDICTION OF BREAST CANCER IN MAMMOGRAM IMAGE USING SUPPORT VECTOR MACHINE

[1]R.JENITA, [2] S.SOLOMON DEVADOSS M.E.,

[1]PG. Scholar Department of Electronics and Communication Engineering, Infant Jesus College of Engineering, Thoothukudi.

[2]Assistant Professor, Department of Electronics and Communication Engineering,,Infant Jesus College of Engineering, Thoothukudi.

ABSTRACT:

Among women, breast cancer is the leading cause of cancer death in recent times. A mammogram is an X-ray image of breast that characterizes normal tissue and different calcification and masses using gray levels showing levels of contrast inside the breast. To analyse an X-ray mammogram image is a challenging task, since the similarities between the cancer growth and other tissue growth. Here the prediction of breast cancer is done by combining one dimensional continuous wavelet transform as the feature selection technique and support vector machine as classifier. In mammogram images segmentation plays an important role to predict and diagnose the breast cancer. Thresholding is the common segmentation method used. In the next step the SVM classifier, that classifies the regions. This method was tested on different images and the highly affected image was predicted.

KEY WORDS:*Mammography, continuous wavelet transform, support vector machine, kernel function, segmentation, thresholding.*

1.INTRODUCTION:

Cancer starts when cell begin to grow out of control. Breast cancer is a tumour that develops from breast tissue. The cells are grouped producing lump or mass called tumour which may be benign or malignant. A malignant tumour is a group of cancer cells that grow or spread to distant areas of the body. The risk factor depends on the stage of detection of the breast cancer. Mammograms are used to detect the early stage and routine check-ups reduce the risk related to death. A uniquely important type of imaging used for the screening of breast cancer is mammography. All women at risk go through mammography screening

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 13, March 2016

procedure for early detection and diagnosis of the breast cancer.A mammogram is an X-ray image of breast that characterizes normal tissue and different calcification and masses using gray levels showing levels of contrast inside the breast. Now-a-days computer aided diagnosis system is used to process many X-ray images in less amount of time. Analysing all the features of the image and to predict the calcification, it is important to extract the available information in the image. Wavelet transform is used to analyse the image containing different structures. To detect the calcifications in mammogram the wavelet functions can be used. One dimensional wavelet transform have been applied to mammographic schemes with varying success. The one dimensional wavelet transform is a less time consuming process.Due to its generalization ability, SVM is used as tool for data classification. Feature of SVM is that it minimizes upper bound of generalization error through minimizing the margin between separating hyper plane and data set. The performance of SVM largely depends on the kernel. The SVM classifier is used to make better classification between the healthy and cancerous tissues.

2. EXPERIMENTAL SETUP:

Analysing the informative features in images with different scale, orientation and representation and finding out the wavelet coefficient is done by using continuous wavelet transform. It performs very well in irregular areas, lines and curves and it is less time consuming. Using SVM classifier the classification and prediction is done. It gives better prediction and finely classifies the location of growth.

3. PROPOSED TECHNIQUE:

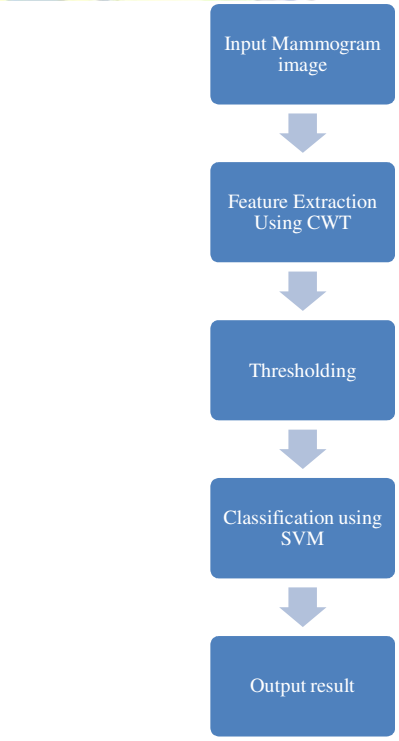


Fig: Flow chart showing the proposed technique

This includes the application of 1D- WT on a set of mammogram images.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 13, March 2016

This results in wavelet coefficients that are sorted in a descending order and select top 100 coefficients for every selected image. Hence the coefficients selected above the threshold to be cancerous and location of growth is to be found and below the threshold as normal image or along with lymph or cysts but not cancerous. For each image the coefficients above and below threshold are selected. During the training phase randomly select the coefficients and apply the SVM classification model on the testing coefficient from the CWT of image set. The classification is done by following these two steps 1. Features are selected from the mammogram image, 2. The selected features with higher wavelet coefficient are classified using SVM classifier.

A. Continuous wavelet transform

Continuous wavelet transform is a process that is used to divide a continuous time function into wavelets. It has the ability to decompose complex information into elementary forms. It possesses the ability to construct a time-frequency representation of a signal that offers very good time and frequency localization. One advantage of using cwt is that it provides image compression that provides significant improvement in the quality of image at high compression ratio. The CWT is used in acoustics processing and pattern recognition. It is very efficient in determining the damping ratio of oscillating signal and very resistant to noisy signal.

B. Threshold

For a data set, the mean is the sum of the values divided by the number of values. The standard average, often simply called the "mean". The mean of a set of numbers Xl, X2 ••• Xn is typically denoted by $\bar{x}$ pronounced "x bar". The mean is often quoted along with the standard deviation: the mean describes the central location of the data, and the standard deviation describes the spread. An alternative measure of dispersion is the mean deviation, equivalent to the average absolute deviation from the mean. It is less sensitive to outliers.

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n} x_i$$

Algorithm for finding threshold:

Input:

Read N, S from CSV file

N: Coefficients of normal images.

S: Coefficients of sick images.

1. Find mean: N & S.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 13, March 2016

2. Find IQR: Mean (N & S)
3. Finding Sigma: IQR (N & S)
4. Finding 1 sl& 3rd quartile of normal and sick.
5. Finding average: avgl, avg2 avgl: Mean (N) +quartile value. Avg2: Mean(S) + quartile value.
6. Threshold= (avgl+avg2) 12;

C. Support Vector Machine (SVM)

A support vector machine is a machine learning method that classifies binary classes by finding and using a class boundary the hyper plane maximizing the margin in the given training data . The training data samples along the hyper planes near the class boundary are called support vectors, and the margin is the distance between the support vectors and the class boundary hyper planes. The SYM are based on the concept of decision planes that define decision boundaries. A decision plane is one that separates between assets of objects having different class memberships. SYM is a useful techniquefor data classification. A classification task usually involves with training and testing data which consists of some data instances. Each instance in the training set contains one "target value" (class labels) and several "attributes" (features).There are number of kernels that can be used in SYM models. These include linear polynomial, RBF and sigmoid.

Polynomial, RBF and sigmoid

$\emptyset = \{x_i * x_j$ linear

$(\gamma x_i x_j +$ coeff) polynomial

exp $(-\gamma |x_i * - x_j|)$ RBF

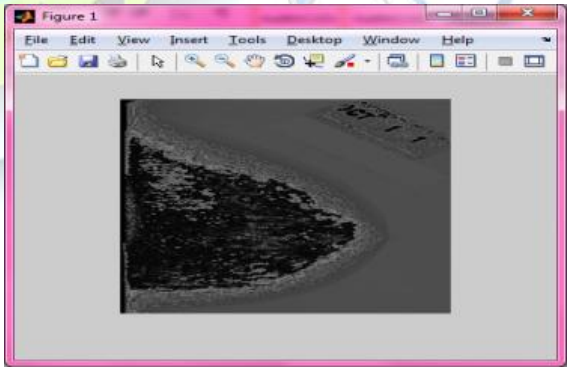
tanh $(\gamma x_i x_j +$ +coeff)$\}$ sigmoid

$$K(x, x') = \exp(-|x-x'|2/(2\sigma 2)).$$

The kernel is then modified in data dependent way by using the obtained support vectors. The modified kernel is used to get the final classifier.

4.RESULT AND DISSCUSSION:

By applying 1D-continuous wavelet transform the suspicious calcification region in the X-ray image can be efficiently reconstructed by the elimination of most normal tissue pixels and background noise.



CWT is constructed to define the feature of calcification tissues. In order to quantify and recognize calcification pixels, the support vector classifier is developed.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 13, March 2016

By support vector machine, the mean square error between output value of the SVM and expected valuecan be minimized so that the examination efficiency for the X-ray images of breast cancer can be greatly improved.

5.CONCLUSION:

The developed method is summarized as, the initial step based on gray level image information is enhanced and the breast cancer is segmented. For every cancerous region, features are extracted to categorize the breast cancer. At last for classification the SVM classifier is used.

References:

1) Medical Imaging and Graphics 35(2011) 1-8 I. Daubechies, "Ten Lectures on Wavelet, "SIAM,pp. 167-213, Philadelphia 1992.

2) Y. Liu, Y. F. Zhung, "FS_SFS: A novel feature selection method for support vector machines", pattern recognition New York, vo1.39, pg.1333-1345, December 2006.

3) Nan-Chyuan Tsai, Hong-Wei Chen, Sheng- Liang Hsu. "Computer-aided diagnosis for early –stage breast cancer by using wavelet Transform".

4) G.Rajakumar, Dr.D.Manimegalai, "FPGA Implementation of DIP based Adulteration Identification in Food Samples", International Journal of Computer Applications (IJCA), Volume 35, No.1, Page No 06 - 11, December 2011, ISSN 0975 – 8887.Impact Factor:0.801.

5) Smola A. J., Scholkopf B., and Muller K. R., "The connection between regularization operators and support vector kernels", Neural Networks New York, voUl, pg 637-649, November 1998.

6) Nandkumar C. Jambhale, Prof. Ajay S. Wadhawe, "Method to Diagnosis and Classify Breast Cancer by using Wavelet Packet & Neural Network", , International Journal of scientific engineering and technology research ISSN 2319- 8885 Vol.(04), Issue 2014 June 2015 pg 2652-2657.

7) GulShairaBanu, AmjathFareeth, prediction of breast cancer in mammogram image using support vector machine and fuzzy C-means, IEEE, jan 2012

8) Y.Ireaneus Anna Rejani, Dr.S.Thamarai Selvi, Early Detection Of Breast Cancer Using SVM Classifier Technique, International Journal on Computer Science and Engineering Vol.1(3), 2009, 127-130

9) A. Mencattini , M. Salmeri , R. Lojacono , F. Caselli, Mammographic Images Enhancement and Denoising for Micro calcification Detection Using Dyadic Wavelet Processing, Instrumentation and Measurement Technology Conference Sorrento, Italy 24-27 April 2006

10) Nalini Singh, Ambarish G Mohapatra, GurukalyanKanungo, Breast Cancer Mass Detection in

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 13, March 2016

Mammograms using K-means and Fuzzy C-means Clustering, International Journal of Computer Applications (0975 – 8887) Volume 22– No.2, May 2011

11) BhagwatiCharan Patel, Dr.G.R.Sinha, Early Detection of Breast Cancer using Self Similar Fractal Method, International Journal of Computer Applications (0975 – 8887) Volume 10– N.4, November 2010