

Comparison Of Different Classification Techniques For Mammogram Image Analysis

A.P.Manjunatha¹, H.S. Sheshadri², M.N.GiriPrasad³

¹Associate Professor, Department of Medical Electronics, Dr.AIT, Bengaluru, India

²Professor, Department of Electronics & Communication, PES College of Engineering, Mandya, India

³Professor, Department of Electronics & Communication, JNTU College of Engineering, Anantpur, India

Abstract

This paper presents experimental work on mammogram image analysis. Texture analysis is carried out using segmentation technique. Here, statistical method have been used to extract features from the segmented tumor area. The obtained features are classified using different classifiers such as K-Nearest Neighbour, Multilayer perceptron, SVM. The method was tested on 100 clinical data. The K-NN classifier achieved an accuracy of 97.6%, Multilayer perceptron achieved an accuracy of 91.59%, SVM will give an accuracy of 96.7%.

Keywords: K-Nearest Neighbour, Multilayer perceptron, SVM, Segmentation, Texture.

1. Introduction

For the early detection of breast cancer Mammography technique is most widely preferred technique by majority of the radiologists. X-Ray Mammography is the most reliable diagnostic procedure for detecting the masses and calcification in the tissue of the breasts. However during the screening procedure, radiologists often miss the detection of a significant proportion of abnormalities in addition for having high rates of false positives. [1] Many computer aided detection techniques are being implemented for the segmentation of these mammogram images and obtaining the desired region of interest, which are having high probability of occurrences masses. [2]

In the digital mammogram images, various segmentation techniques are implemented for segmenting the lesions. Radial gradient index based algorithm and a probabilistic algorithm are two such techniques. These segmentation techniques begin with a point, which is known as seed point. This point will be present in the suspected lesion. [3] The segmentation techniques are generally

Consisting of gray- level thresholding, region growing or active contours methods. [4]

In this paper we have made an attempt for segmentation of these mammogram images using texture analysis [5]. This procedure is using statistical approach for finding the suspected region of interest, having high probability of the occurrence of masses or calcification. From this region of interest 6 features are obtained. Further classification of these features is performed using Artificial neural Networks (ANN), K-Nearest Neighbour (KNN) and

Support vector Machine (SVM) algorithms, with respect to the clinical diagnosis (Malignant or Benign). Comparison of the results of accuracy obtained from these classifier algorithms is also performed. The dataset of real time mammogram images are obtained from Methodology.

1.1 Texture

To quantify the texture contents of a region we will use statistical approaches for computing texture based on statistical measures.

Statistical Approaches

The texture analysis is based on the statistical properties of the intensity of histogram. [6] It is based upon Statistical Moments [paper on statistical moments]. The expression for the n^{th} order moments about the mean is given by equation (1)

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i) \quad (1)$$

Where z_i is a random variable indicating intensity, $p(z_i)$ is the histogram of the intensity levels in a

region, L is the number of possible intensity levels and

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \quad (2)$$

is the mean(average) intensity given by equation (2). Mean is the measure of average intensity.

Standard deviation is the measure of average contrast given by the equation (3).

$$\sigma = \sqrt{\mu_2(z)} = \sqrt{\sigma^2} \quad (3)$$

Smoothness measures the relative smoothness of the intensity in a region. R is 0 for a region of constant intensity and approaches 1 for regions with large excursions in the values of its intensity levels.

The variance used in this measure is normalized to the range [0, 1] by dividing it by $(L-1)^2$ given in the equation (4)

$$R = 1 - 1 / (1 + \sigma^2) \quad (4)$$

Variance is the measure of contrast.

Third moment measures the skewness if a histogram. This measure is 0 for symmetric histograms, positive by histograms skewed to the right (about the mean) and negative for histogram skewed to the left. Values of this measure are brought into a range of values comparable to the other five measures by dividing μ_3 by $(L-1)^2$ also, which is the same divisor we use to normalize the variance given by equation (5).

$$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i) \quad (5)$$

Uniformity measure is maximum when all the gray levels are equal (maximally uniform) and decrease from there, given in equation (6)

$$U = \sum_{i=0}^{L-1} p^2(z_i) \quad (6)$$

Entropy is the measure of the randomness given by equation (7)

$$e = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \quad (7)$$

1.1 Classification of the obtained features

1.1.1 ANN using Multilayer Perceptron

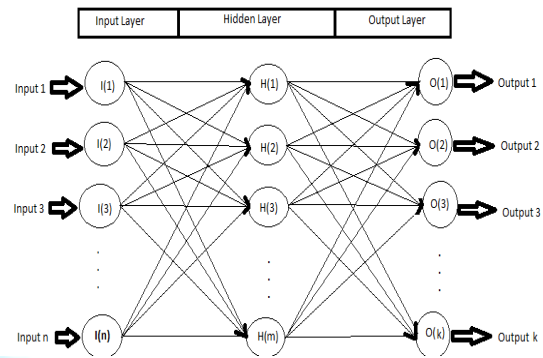


Figure 1. Multilayer Perceptron

It is the type of classification technique which is implemented by artificial neural networks. This will make use of neurons (also called perceptrons in this case). Various concepts for designing MLPs, e.g., fully parallelized, serialized, or even a mixed architecture can be used. Three layer MLP is shown in Figure 1. It consists of several artificial neurons. These neurons are represented by perceptrons. Figure 1 represents three layers MLP. Each layer contains a set of neurons. The first layer is the input layer and represents the input values. The hidden layer is the second layer. It is meshed with the previous layer. The output layer generates the outputs and receives the values from the hidden layer. The determination of weight factors are known as training, where each neuron of a layer is connected to each neuron of the following layer. The MLP is a “feed forward”, fully connected ANN [7,8]. Hence, the input value of a neuron is the sum of the n neuron outputs of the previous layer as shown by equation 8

$$Neuron_{Input} = \sum_{i=1}^n Input_i * WeightFactors_i \quad (8)$$

1.1.2 Support Vector Machine (SVM)

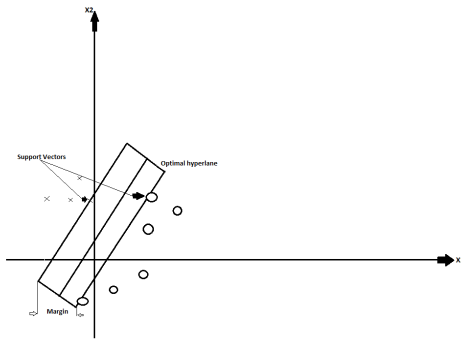


Figure 2. Hyperplane separating 2 classes in a SVM classifier

The basic idea of an SVM classifier is illustrated in Figure. 2. This figure shows the simplest case in which the data vectors (marked by 'Xs and 'Os) can be separated by a hyperplane. In such a case there may exist many separating hyperplanes. Among them, the SVM classifier seeks the separating hyperplane that produces the largest separation margin. Such a scheme is known to be associated with structural risk minimization. In the more general case in which the data points are not linearly separable in the input space, a nonlinear transformation is used to map the data vector x into a High-dimensional space (called feature space) prior to applying the linear maximum-margin classifier. To avoid the potential pitfall of over-fitting in this higher dimensional space, an SVM uses a kernel function in which the nonlinear mapping is implicitly embedded. According to Cover's theorem, a function qualifies as a kernel provided that it satisfies the Mercer's conditions. With the use of a kernel, the discriminate function in an SVM classifier has the following form:

$$g(x) = \sum_{i=1}^{L_5} \alpha_i d_i K(x_i, x) + \alpha_0 \quad (9)$$

where $K(\cdot, \cdot)$ is the kernel function, x_i are so-called support vectors determined from training data, L is the number of support vectors, d_i is the class indicator (e.g., +1 for class 1 and -1 for class 2) associated with each x_i , and α_i are constants, also determined from training. By definition, support vectors (Fig. 2) are elements of the training set that lie either exactly *on* or *inside* the decision boundaries of the classifier. In essence, they consist of those training examples that are most difficult to classify. The SVM classifier uses these "borderline" examples to define its decision boundary between the two classes. This in philosophy is quite different from a classifier that is based on minimizing learning error alone. Note that in a typical SVM learning

problem only a small portion of the training examples will typically qualify as support vectors.

1.1.3 K Nearest Neighbour (KNN) Algorithm

K-nn algorithm is among the instance-based classifiers. In instance-based methods, system parameters or classifying system units simply consist of the samples that are presented to the system. This algorithm assumes that all instances correspond to points in the n -dimensional space R^N . Nearest neighbors of a sample in this space are determined by standard Euclidean distance. Let x be a sample and it is defined by a feature vector of: $\{a_1(x), a_2(x), \dots, a_n(x)\}$ (10)

here $a_r(x)$ is the r th feature of x sample. The $d(x_i, x_j)$ Euclidean distance between x_i and x_j samples is defined by

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (11)$$

k -nn algorithm uses an $f(\cdot)$ function. For classification applications, this function is the class of presented sample. If we denote this by $f(x_i)$, the procedure of k -nn algorithm can be summarized as follows: k -nn algorithm stores all training data and corresponding classes of this training data as system units. Let $\{x_i, f(x_i)\}$ be a vector indicating individual training sample and the corresponding class of this sample. During the classification in the system, k -nearest system units to presented x_i sample are determined via Eq. (11). The class of presented sample is approximated according to the number of these k -nearest units. The class of nearest samples that have the highest percentage in k -nearest units be this class estimation; $\hat{f}(x_i)$.

2. Results and Discussion:-

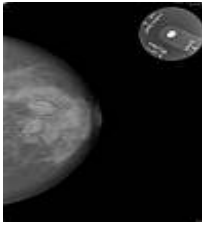


Fig:-3

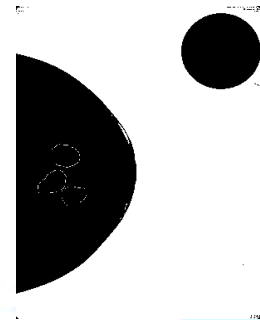


Fig:-4

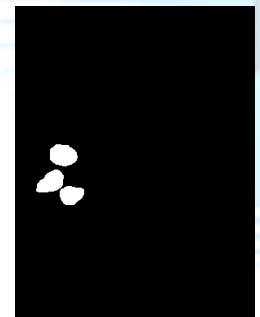


Fig:-5

-----Region 1 -----

Avg Intensity = 141.399823,
 Avg Contrast = 16.011316
 measure of smoothness = 0.003927,
 3rd Moment = -0.177547,
 Uniformity = 0.025809,
 Entropy = 5.668142

-----Region 2 -----

Avg Intensity = 143.552424,
 Avg Contrast = 14.360004,...
 measure of smoothness = 0.003161,
 3rd Moment = -0.173961,
 Uniformity = 0.029514,
 Entropy = 5.437631

Region 3 -----

Avg Intensity = 130.912472,
 Avg Contrast = 14.654976,...
 measure of smoothness = 0.003292,
 3rd Moment = -0.138656,
 Uniformity = 0.027146,
 Entropy = 5.555134

Fig:-3 shows Original Image.

Fig:-4 shows boundary of ROI

Fig:-5 Segmented Region

The confusion matrix obtained by ANN using MLP

Table 1

	YES	NO
YES	TP	FN
NO	FP	TN

Table 2

	YES	NO
YES	73	14
NO	1	125

Training ratio 80% Efficiency 91.5% . The sensitivity is achieved 84% and specificity is about 99.21%.

In SVM classifier

Table 3

	YES	NO
YES	82	5
NO	1	125

Training ratio 80% Efficiency 96.7%. The sensitivity is achieved 94.25% and specificity is 99.20%

In K-NN classifier

Table 4

	YES	NO
YES	86	1
NO	3	123

Training ratio 80% Efficiency 97.6 %

The sensitivity is achieved 98.85% and specificity is about 97.21%.

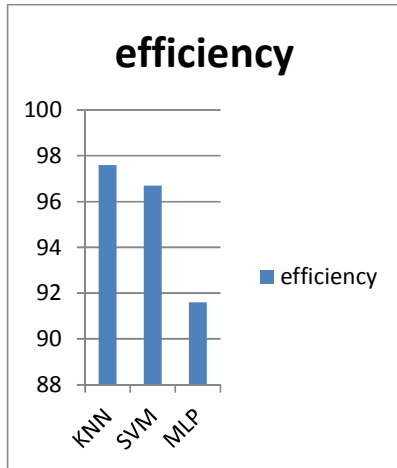


Fig:-6

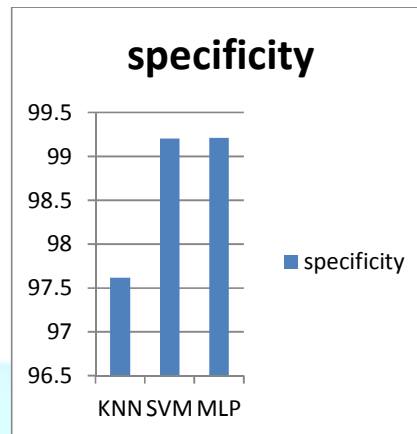
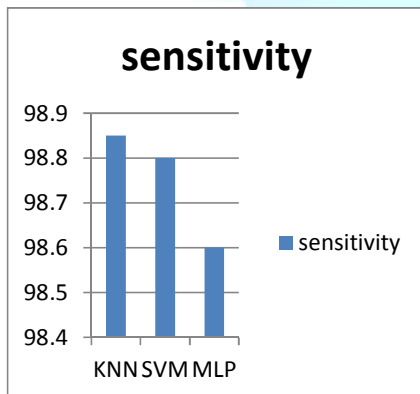


Fig:- 7

Conclusion:

The mammogram images are segmented using texture analysis by using statistical approaches. The desired region of interest is obtained from the marked calcifications on the images. Further 6 features are obtained from the region of interest. These features are classified using KNN, Multilayer perceptron and SVM algorithms. Further, respective efficiency, sensitivity and specificity are obtained from the respective classifiers and the obtained results are compared.

Acknowledgement:

The authors are grateful to Dr.Umashankar, Radiologist, Padmashree Diagnostic Center, Bangalore for providing Medical image database used in this study and for his valuable guidance to our work. Also we are thankful to staff of Padmashree Diagnostic Center.

References:

- [1] Mohammed J. Islam, Majid Ahmadi and Maher A. sid-Ahmed "An efficient automatic Mass classification method in Digitized Mammograms using Artificial Neural Network (IAIA) vol.1 no.3 pp.1-13, 2010.
- [2] R Mckenna "The Abnormal Mammogram Radiographic Findings, Diagnostic options, Pathology, and stage of cancer Diagnosis" PP 224 -225, 1994 Wiley online Library.

[3]Matthew a. Kupinski and Maryellen L Giger, Member, IEEE “Automated Seeded Lesion Segmentation on Digital Mammograms “Vol. 17, no.4 pp 510-517,1998 IEEE Transactions on Medical Imaging.

[4]R. Campanini, E. Angelini, E. Iampieri and M. Masotti, “ A fast algorithm for intra-breast segmentation of digital mammograms for CAD systems,” IWDM 2004 .

[5]H.S.Sheshadri and A.Kandaswamy “Breast Tissue classification using statistical feature Extraction of Mammograms” Medical Imaging and Information Sciences Vol.23 No.3 September 2006

[6]J. Skodzik, V. Altmann, B. Wagner, P. Daniels and D. Timmermann,”Highly Integrable FPGA- Based Runtime-Configurable Multilayer Perceptron,” 2013

[7]Z.Liu, “Parameter selection for Gaussian Radial Basis Function in Support Vector Machine classification “ PP-576-581,2012.

[8]EI- Naqa “Support vector machine Learning for detection of Microcalcification in Mammograms” pp 201-204, Biomedical Imaging, 2002 IEEE International Symposium

[9] R. C. Gonzalez R. E. Woods, A.H. Scoott, and P.P. Hall, Digital Image Processing 2003.

