International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 3 (2015) pp. 7751-7756

http://www.ripublication.com

BPN Based Classification of Digital Mammograms Using Boundary and Texture Descriptors

[1]V.Vedanarayanan, [2]N.M.Nandhitha

*[1]Asst. Prof., Dept. of ETCE, Jeppiaar Nagar,
Old Mamallapuram Road, Chennai 600 119*
[2]Prof., Dept. of ECE, Jeppiaar Nagar, Old Mamallapuram Road, Chennai 600 119
Email:veda_gopal@rediffmail.com, nandhi_n_m@yahoo.co.in

Abstract

Breast cancer causes higher mortality rate in women. Reliable detection of breast cancer is the only method for reducing the mortality rate. Digital mammography is the golden standard for determining breast cancer. Digitized mammographs are then interpreted by certified radiologists. However manual interpretation is subjective in nature and is affected by operator fatigue. Hence in recent years the paradigm has shifted to Computer Aided Analysis. Automated classification algorithms are developed to classify mammographs as normal or benign or malignant. It involves extracting the features and training the classifier with the extracted features. The efficiency of the proposed algorithm is dependent on the effective representation of the features and the choice of the classifier. In this paper, an efficient set of exemplars is created with both regional and texture features. These features are then used to train and test the ANN based classifiers. It is found that the proposed technique has higher accuracy and sensitivity than the classifiers developed with the other features cited in the literature.

Keywords: Mammographs, Regional descriptors, Haralicks features, Back Propagation Network.

Introduction

According to American Cancer Society, breast cancer is the most commonly occurring disease in women. Breast cancer is due to the uncontrolled growth of cells in the breast area or under the arm pits of women. These cells form a mass or lump which the physicians refer as tumors. These tumors can be classified into two types namely Malignant and benignant. Malignant tissues are cancer tissues whereas benignant is non cancer tissue growth with a chance of becoming malignant if left untreated. Benignant growth is due to calcification which is common in lactating mothers. Usually architectural distortations, micro calcification masses are analyzed to detect the presence of malignant tumors. It leads to higher mortality rate in women. However if breast cancer is detected, then the mortality rate can be reduced. Conventional diagnostic techniques include MRI, PET, CT and Digital mammography. Of the various diagnostic techniques, Digital mammography is the most widely accepted, reliable technique which captures the X-ray of the region under observation. Mammography is the technique used to detect the unwanted tissue growth in the breast regions. A low dose X-Ray is passed through the test area and detected by a high resolution, high contrast detectors. The image thus recorded through this procedure is known as mammogram, which can be used for both screening and diagnosis. In a mammogram, tumor regions correspond to higher intensities. However the contrast of the mammographs is poor due to the virtue of imaging equipment. Hence manual interpretation of mammographs is strongly dependent on the expertise of the radiologist and is also affected by operator fatigue.

In recent years, extensive research is carried out in developing computer aided analysis and classification of breast cancer based on digital mammographs. Automated classification system involves the following steps: Preprocessing mammographs, feature extraction and classification. However a standard is not yet evolved due to the inability to determine the most suitable features for extracting the information and the suitable classifier. Hence it necessitates identifying suitable feature vectors for describing the mammographs and optimizing the classifier parameters. In this paper, feature extraction is performed after extracting the tumor region from the mammographs. Back Propagation Network based classifier is used for classifying the mammographs. This paper is organized as follows: Section 2 provides the related work. Section 3 describes the proposed feature extraction and classification technique. Section 4 provides the Results and Discussion and Section 5 concludes the work and provides the feature work.

Related Work:

In recent years, extensive research is carried out in developing automated breast cancer classifiers using soft computing and Support Vector Machines. Mumtaz et al (2008), compared the performance of Self organizing maps (SOM), Adaptive Resonance Theory (ART) and BPN for classification from breast mammographs. In this paper the authors had considered both supervised and unsupervised trainings. In SOM, Euclidean distance between the weights were calculated to get the best match. For Back Propagation and Resonance Theory networks, set of features were extracted and used to train the networks. To measure the performance of these networks, Total run time, Speed and Error were determined. It was observed that the classical BPN method was superior when compared with the other methods.

F.Paulin et al (2011) made a study on BPN based classification. Performance of six different BPN Training algorithms such as Batch training, Batch Descent training, Batch gradient with moment, Conjugate Gradient, Quasi Newton Algorithm and

Levenberg Marquardt algorithm(LM) were compared on Wisconsin Breast Canser Data Set(WBCD). Learning rate considered was 0.7 and maximum epochs was fixed as 1000.It was observed that LM based BPN was superior when compared with other training algorithms and an accuracy of 98.29% was achieved.

Baljit Singh and Pratap Sing (2011) used Resilient BPN algorithm for classification of micro calcification .Automated Computer Aided System (CAD) was proposed for identification of Ductal carcinoma in situ (DCIS), which was the most common type of breast cancer. DCIS can be detected by properly identifying micro calcification in the breast. Standard Deviation(SD), relative Smoothness(R), Skewnwss, Kurtosis, Busyness(B),Potential of Point(P), Mean Energy(E) and Point mask (PM) were some of the features used in this paper for identification of micro calcification. Efficiency was calculated by considering MSE, number of Epoch, Area under Receiver Operating Characteristic. It is observed that Resilient Back propagation was best in terms of speed and accuracy.

P.valarmathi et al (2013) had presented a method for breast tumor prediction based on extracting the mammogram features by Gabor filter and classifying these with Artificial Neural Network. Back Propagation network with different activation function such as sigmoidal function, Tanh function, Gaussian Activation are used for classification. Learning rate is 0.1 and momentum is fixed as 0.5. A total of 10 neurons was considered for the single hidden layer and 25 neurons for input layer and two neurons for output layer is considered in this work. Precision, Recall and F-measure were used to evaluate the efficiency. It was observed that Gaussian function provided better results than the other techniques.

Po-Tsun Liu et al(2013) Presented a Fractal based BPN classifier for classification of breast calcifications. Gabor wavelet was used for feature extraction. Gabor smoothening filter was used to remove noises. Cubic curve method was used for contrast adjustments. Gabor wavelet and fractal dimensions were used to extract features from the breast region. BPN network was used for classification. It was observed that the Gabor wavelet combined with fractal dimension the efficiency of the proposed system was 89.55%.

Yasmin M George et al (2012) proposed a Computer aided system for breast tumor detection. Hough transform in conjuction with watershed algorithm was used for segmenting the Region of Interest. Shape based and texture based features were extracted. A clustering step is performed for classification of tumors. BPN, SVM, PNN networks were trained on these feature values and performance were analyzed based on error rate, correct rate, specificity & sensitivity. It was observed that 99% of sensitivity is obtained.

Bpn Based Classifier For Mammograph Classification

The proposed methodology involves the following steps: Mammograph acquisition, feature extraction and classification. In order to perform the research work, a total of 25 images were collected from leading hospitals and scan centres in Chennai. In the considered dataset, two mammographs are that of normal persons and the remaining are that of abnormal cases. The obtained mammographs are gray scale images and the spatial resolution is 335x448. Eight bits are used for representing the intensity of the pixels. From the manual interpretation of mammographs, it is evident that higher intensity regions correspond to tumors while the lower intensity regions indicate the background. Having acquired the mammographs, the next step is to extract the features. In order to extract the features, initially the breast region is isolated. Later segmentation by thresholding is used to extract the tumor region. Threshold is 150 and all pixels greater than the threshold are retained and the remaining pixels are discarded as background pixels. After extracting the tumor regions, the regional descriptors namely Area, Major axis length, Minor axis length, Equivdiameter, Eccentricity, Orientation, Perimeter, Solidity are extracted. In order to improve the classifier performance, Haralick's features are also obtained on the original mammographs. As a total 12 features are used as input parameters for generating the exemplars. The output parameter is coded as 1 for normal and 2 for tumor regions. A Back Propagation Network based classifier is chosen for classifying the mammographs. It is a five layered network with three hidden layers and one output layer. The number of neurons in the first, second and third hidden layer are 24, 12 and 6 as it is found that the 2x, x,x/2 is the optimal choice (ref). The activation functions are tansigmoidal for hidden layers and purelinear for output layer.

Results and Discussion

The exemplars generated for training the neural network based classifier is as shown in Table 1. Two sets of exemplars are used for training and testing the neural network based classifier. Performance of the neural network is measured in terms of sensitivity, accuracy and specificity as shown in Table 3. The sensitivity and accuracy of the neural network for the trained dataset is 100% which indicates that the neural network has performed well for the trained dataset. The performance of the neural network for the test dataset is shown in Table 2 for the neural network trained with only regional descriptors and with both regional descriptors and Haralick features. It is found that the accuracy in the first case is 66.67% and 83.33% respectively. Sensitivity is 100% for both the cases. However specificity is 0% and 50% respectively. Hence it is found that neural network based classifier trained with 12 input vectors provides better results than that of 8 input training.

Table 1: Exemplar generation for training BPN

Area	Orientation	Perimeter	Solidity	Major axis length	Minor axis length	Equiv diameter	Eccentricity	Contrast	Correlation	Energy	Homogenity
14618	55.8748	519.1026	0.9643	180.1815	106.0184	136.4266	0.8086	0.1777	0.983	0.3119	0.9526
2033	-89.8964	671.6569	0.9356	385.2063	7.1799	50.8773	0.9998	0.1788	0.9697	0.3168	0.9556
27	-1.321	25.0711	0.9348	11.0404	5.2228	7.3993	0.881	0.1372	0.9762	0.5267	0.9646
2560	-89.8893	790.2426	0.9671	7.669	445.8979	57.209	0.9999	0.1283	0.9838	0.4236	0.9686
1896	90	640	1	364.8854	6.9282	49.1331	0.9988	0.1122	0.9844	0.5184	0.9755
28466	66.1081	945.7788	0.8014	293.29	139.8854	190.3787	0.8789	0.1893	0.9848	0.2664	0.9453
7687	86.8799	23.0711	0.8974	147.1874	69.6756	98.9313	0.8809	0.1268	0.9803	0.3809	0.9623
6971	77.5777	680.5168	0.5951	167.8164	70.8195	94.2112	0.9066	0.1509	0.9815	0.4732	0.9597
2863	-89.9633	827.0711	0.9359	471.0182	8.1446	60.3762	0.9999	0.1766	0.9788	0.3922	0.9561
49529	86.5093	1098	0.9501	328.5138	199.5242	251.1221	0.7944	0.2068	0.9807	0.2289	0.9354
352	2.6112	119.598	0.6365	36.2286	18.5412	21.1703	0.8591	0.0299	0.9937	0.3078	0.985

Table 2: Relationship between desired and actual Outputs for network trained with 8 inputs and 12 inputs.

target	8 inputs	12 inputs
2	2	2
1	2	2
2	2	2
2	2	2
2	2	2
1	2	1

Table 3: Performance metrics of classifier trained with 8 input and 12 input exemplars.

Parameters	8 input parameters	12 input parameters
Sensitivity	100%	100%
Specificity	0%	50%
Accuracy	66.67%	83.33%

Conclusion

In this paper, BPN based classifier has been successfully developed to classify mammographs as normal or abnormal. The developed neural network classifier is trained with two sets of input parameters and the performance is measured in terms of sensitivity, accuracy and specificity. It is found that inclusion of Haralick features has increased all the above performance metrics. It is because Haralick's metrics captures the texture information in the mammographs. However in this case, the research database is limited to 25 mammographs. Hence in order to improve the performance of the classifier it is necessary to collect a very large database. Also mammograph preprocessing and statistical parameters must also be included.

References

1. F.Paulin, A.Santhakumaran, Jan 2011, " Classification of Breast cancer by comparing Back propagation training algorithms", International Journal on Computer Science and Engineering (IJCSE)vol-3,no 1, ,PP 327 -332.
2. Baljit Singh Khehra, Amar Partap Singh , October 2011, " Classification of Clustered Microcalcifications using esilient Backpropagation Training Algorithm", special issues on IP Multimedia Communications- IJCA journal, pp 118-124.
3. P. Valarmathi & Dr. V. Radhakrishna , 2013, "Tumor Prediction in Mammogram using Neural Network Gobal journal of ",Global journal of Computer science and technology,Neural & Artificial Intelligence,Vol 13 issue 2.Pp 19-24.
4. Po-Tsun Liu, Chiun-Li Chin, Jing-Shian Chan, Hao-Hung Tsai,2013, "The Classification of Breast Calcification using Texture Feature of Fractal-based and Gabor Wavelet-based " , Proceedings in 3rd International Conference on Multimedia technology ,ICMT13, pp 1830-1837.
5. Yasmin.M.George, 2012, "Breast Fine Needle Tumor Classification Using Neural Networks", International Journal of computerscience issues,Vol - 9,issue 5,No 2,Pp 247-256.
6. K. Mumtaz, S.A.Sheriff, Dr. K. Duraiswamy,October 2008, " Evaluation of Neural Network based Classification of Neural Network Based Classification system for Clinical Cancer Data Classification",Journal of Computer Applications,Vol – 1,No 4,Pp 9 – 14.
7. Xiaoming Liu, Jinshan Tang,2014, "Mass Classification in Mammograms Using Selected Geometry and Texture eatures, and a New SVM-Based Feature Selection Method", IEEE Systems Journal,Vol 8,No 3,Pp 910-920.
8. Liyang Wei,Yongyi Yang, Robert M. Nishikawa, and Yulei Jiang, 2005 , "A Study on Several Machine-Learning Methods for Classification of Malignant and Benign Clustered Microcalcifications", IEEE transactions on medical imaging, Vol. 24, No. 3, Pp 371-380.
9. Jae Young Choi, Dae Hoe Kim, Konstantinos N. Plataniotis,and Yong Man Ro, 2012, "Combining Multiple Feature Representations and AdaBoost Ensemble Learning for Reducing False-Positive Detections in Computer-aided Detection of Masses on Mammograms", IEEE 34th Annual International Conference of the IEEE EMBS San Diego, California USA, Pp 4394-4397.