Effects of Dimension Reduction In Mammograms Classification

Canan ORAL[1], Hatice SEZGİN[2]

[1]Amasya University, Faculty of Technology, Department of Electrical and Electronics Engineering, Amasya, Turkey
canan.oral@amasya.edu.tr
[2] Ondokuz Mayıs University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Samsun, Turkey
hsezgin@omu.edu.tr

Abstract

Breast cancer is the most common type of cancer among women and causing deaths in women. In this paper, a CAD system is presented to investigate effects of dimension reduction for classifying mammograms. Proposed system consists of preprocessing, feature extraction, dimension reduction and classification steps. Multiscale top-hat transform is used to enhance mammograms and to remove noise. First order and second order textural features are extracted from enhanced mammograms. Principal component analysis (PCA) is used for dimension reduction. Two multilayer perceptron neural networks (MLP) are used to classify mammograms as normal or abnormal. All twenty features (without PCA) and selected seven features by PCA are applied each of two classifiers. First MLP classifier with all features achieved accuracy of 79,4%. Second MLP classifier with selected features by PCA achieved accuracy of 91,1%. PCA feature dimension reduction improved the classification performance, increasing accuracy value from 79,4% to 91,1%.

1. Introduction

Breast cancer is the most common type of cancer among women and causing deaths in women [1-5]. The early detection of breast cancer can be important key to survival [2].

Mammography is x-ray imaging technique of different views of breast, and plays a major role in the early detection of breast cancer [1-3].

A computer aided diagnosis (CAD) technology is important development to mammographic imaging. A computer aided diagnosis (CAD) system uses computerized algorithms in order to detect breast abnormalities. CAD system works like a second opinion in helping radiologists during diagnostic process. CAD system helps radiologists interpret mammography and avoid overlooking a cancer [1,2, 6-8]. Several techniques have been presented in literature for classification of mammograms.

Yu and Guan proposed a CAD system based on feature extraction and neural neural network [1,9]. Verma and Zakos presented a CAD system based on fuzzy-neural and feature extraction for detecting and diagnosis microcalcifications on digital mammograms [1,10]. El-Naqa et al. employed support vector machines to detect microcalcifications in mammograms [11]. Zhang et al. developed new mixed feature two stage false positive reduction scheme to detect microcalcifications on mammograms [12]. Masotti used ranklet transform and support vector machines to classify mammograms [1, 13]. Zhang et al. proposed a new algorithm based on twin support vector machines for microcalcification detection [14].

2. Material and Method

In this paper a CAD system is presented to investigate effect of dimension reduction for classifying mammograms. Proposed system consists of preprocessing, feature extraction, dimension reduction (feature selection) and classification steps. In preprocessing step, multiscale top-hat transform is used to enhance mammograms and to remove noise. In feature extraction step, first order and second order textural features are extracted from enhanced mammograms. In feature selection step, principal component analysis is used for dimension reduction. In classification step, multi layer neural networks are used.

MIAS (Mammographic Image Analysis Society) database is used. Every image is 1024x1024 pixels in database [15]. Mamographic images are cropped that contain lesion to reduce complexity. The dimensions of our images are 300x300 pixels.

2.1. Multiscale Top-hat Transform

The top-hat transform is used to detect bright region smaller size structuring element. This transform enhances bright objects that differ from background in images. Top-hat transform is difference between the original image and its opening. In multiscale top-hat transform with different scales structuring element extract scale-specific bright features from image. This transform is local contrast enhancement techniques [1, 16, 17]. Multiscale top-hat transform is defined as ;

$$top(g) = g - (g \circ nB) \qquad (1)$$

Where g is image, B is structuring element of defined shape and n is an integer representing the scale factor.
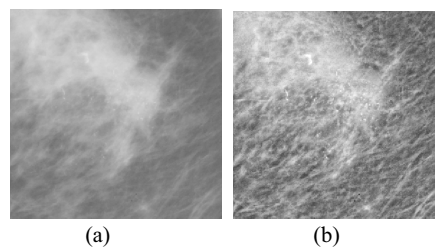

Fig. 1. (a) Original mammographic image b) enhanced mammographic image.

Fig. 1 shows an example mammogram processed by multiscale top-hat transform. As a scale factor n=4 and structuring element disk shape with radius of 6 pixels are

chosen. Fig. 1a is original mammographic image and Fig. 1b is enhanced mammographic image.

2.2. Feature Extraction

It is necessary extract features to classify mammographic abnormalities. In this paper, first order and second order textural features are extracted from enhanced mammograms. [1, 18]

a) First order textural features

First order textural features are based on statistical moments of gray-level histogram of an image or region. Considering the gray level histogram $\{h_i, i=0,1,\ldots\ldots, L-1\}$, if N is total number of pixels and L is number of distinct gray levels in image, then the normalized histogram of image is the set $\{p_i, i=0,1,2\ldots\ldots, L-1\}$, where $p_i=h_i/N$ [1, 18, 19].

In this paper six first order textural features are extracted. These features are mean, standard deviation, skewness, kurtosis, uniformity and entropy.
They are described as;

Mean;
$$m = \sum_{i=0}^{L-1} i p_i \tag{2}$$

Standard deviation (σ) ;
$$\mu_2 = \sum_{i=0}^{L-1}(i-m)^2 p_i \quad \sigma = \sqrt{\mu_2(z)} \tag{3}$$

Skewness;
$$\mu_3 = \frac{1}{\sigma^3}\sum_{i=0}^{L-1}(i-m)^3 p_i \tag{4}$$

Kurtosis;
$$\mu_4 = \frac{1}{\sigma^4}\sum_{i=0}^{L-1}(i-m)^4 p_i \tag{5}$$

Uniformity;
$$U = \sum_{i=1}^{L-1} p_i^2 \tag{6}$$

Entropy;
$$e = -\sum_{i=0}^{L-1} p_i \log_2 p_i \tag{7}$$

b) Second order textural features

Second order textural features are computed gray-level co-occurrence matrix (GLCM) [1, 20,21].

Gray-level co-occurrence matrix provide valuable information about relative position of neighboring pixels in an image. Gray-level co-occurrence matrix is calculated from angular matrices [20,21].
Angular matrices are defined as;

$$P(i,j,d,0^0) = \# \begin{cases} ((k,l),(m,n)) \in (L_y x L_x) x(L_y x L_x) \\ |k-m|=0, \quad |l-n|=d, \quad I(k,l)=i, \quad I(m,n)=j \end{cases} \tag{8}$$

$$P(i,j,d,45^0) = \# \begin{cases} ((k,l),(m,n)) \in (L_y x L_x) x(L_y x L_x) \\ (k-m=d, \quad l-n=-d) \, veya \, (k-m=-d, \quad l-n=d), \\ I(k,l)=i, \quad I(m,n)=j \end{cases} \tag{9}$$

$$P(i,j,d,90^0) = \# \begin{cases} ((k,l),(m,n)) \in (L_y x L_x) x(L_y x L_x) \\ |k-m|=d, \quad l-n=0, \quad I(k,l)=i, \quad I(m,n)=j \end{cases} \tag{10}$$

$$P(i,j,d,135^0) = \# \begin{cases} ((k,l),(m,n)) \in (L_y x L_x) x(L_y x L_x) \\ (k-m=d, \quad l-n=d) \, veya \, (k-m=-d, \quad l-n=-d), \\ I(k,l)=i, \quad I(m,n)=j \end{cases} \tag{11}$$

Where $P_{\theta,d}$ is angular matrix at θ^o (0°, 45°, 90°, 135°) and calculated for distance of d between two pixels whose coordinates are (k,l) and (m,n) , respectively; i and j are gray levels; and I is intensity function. Textural features are calculated for each GLCM.

In this paper GLCM is calculated for parameter values $d = 1$ and $\theta = 0^o$. Fourteen features extracted for calculated GLCM. These features are contrast, correlation, entropy, uniformity, difference variance, difference entropy, normalized inverse difference moment, energy, variance, sum average, sum variance, sum entropy, information measurement of correlation 1,2 [20].

2.3. Dimension Reduction (Feature Selection)

Dimension reduction plays an important role in classification performance. Feature vector has usually high dimension. The aim of feature selection is to reduce dimensionality of the measurement space to a space suitable for application of classification algorithms. The feature space can be transformed space that has lower dimension than the original [22, 23].

In feature selection step, principal component analysis (PCA) is used for dimension reduction.

Principal component analysis is approach to reduce dimensionality. PCA transforms a number of correlated variables in to a smaller number of uncorrelated variables called principal components [24]. First covariance matrix is calculated using feature vector. Next, eigenvectors and eigenvalues are computed and sorted according to decreasing eigenvalue. Bigger variances of data distribution are more effective to discriminate the classes than smaller variances[25].

In this study we reduce the feature vector dimensionality from twenty to seven using PCA.

2.4. Classification

In classification step, multilayer perceptron neural networks (MLP) is used for classification mammograms into normal or abnormal. Multilayer perceptron consists of an input layer, one or more hidden layer and an output layer. MLP passes weights assigned to different layers, and determines the output and compares it with target output. Then it propagates error signal and adjust the connection weights correspondingly [1, 24,26,27]. Multi layer neural network architecture is shown in Fig. 2.
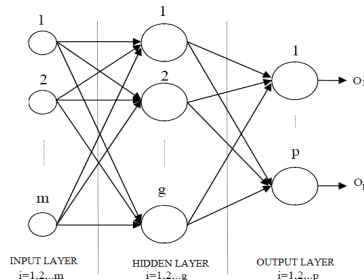

Fig. 2. Multilayer perceptron neural network architecture

For evaluation of the classification performance; sensitivity, specificity and accuracy are calculated. The following expressions are used;

$$Sensitivity = \frac{TP}{TP + FN} x100 \qquad (12)$$

$$Specificity = \frac{TN}{TN + FP} x100 \qquad (13)$$

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} x100 \qquad (14)$$

Where TP,TN,FP, and FN are respectively true positives, true negatives, false positives, and false negatives [1].

3. Experimental Results and Discussion

The classification process is performed with and without PCA using MLP classifier. MLP classifier is applied to training and testing to classify mammograms into normal and abnormal. Two MLP classifiers are constructed and classification results are compared. Raw feature vector and PCA feature vector are applied two constructed MLP classifiers. Each of two MLP have three layers.

First MLP classifier used all twenty features (without PCA) to classification. In MLP; input layer has twenty neurons, hidden layer has thirty neurons and output layer one neuron. Classification results are given in Table 1.

Table 1. First MLP Classification results without PCA

Classifier	Feature number	Sensitivity (%)	Specificity (%)	Accuracy (%)
MLP	20	75	80,77	79,4

Second MLP classifier used selected seven features by PCA to classification. In MLP ; input layer has seven neurons, hidden layer has eleven neurons and output layer one neuron. Classification results are show in Table 2.

Table 2. Second MLP Classification results with PCA

Classifier	Feature number	Sensitivity (%)	Specificity (%)	Accuracy (%)
MLP	7	87,5	92,3	91,1

The second MLP with PCA achieved best classification rate.

4. Conclusions

In this paper, we implement a CAD system and investigate effects of dimension reduction for classifying mammograms. First, mammograms are enhanced using multiscale top-hat transform. Next, six first order and fourteen second order textural features are extracted from enhanced mammograms. Next, we apply principal component analysis to twenty dimensional raw feature vector in order to reduce dimension. Next, we select best seven PCA features. Finally, two MLP architecture are employed for classification and their classification results are compared. Raw feature vector and selected features with PCA are applied each of two classifiers. First MLP classifier used all twenty features achieved accuracy

of 79,4%. Second MLP classifier used selected seven features achieved accuracy of 91,1%. According to classification rates, second MLP used features selected by PCA method has best classification performance.

Experimental results demonstrate that, seven features selected with PCA are more distinguishable than twenty raw features. We observe that, dimension reduction with PCA increase performance for classifying mammograms as normal or abnormal.

It can be investigated and compared classification performance using different feature selection methods and different classifiers.

5. References

[1] C. Oral, H. Sezgin, "Çok Katmanlı Sinir Ağı Kullanarak Mamogramların Sınıflandırılması", *Elektrik-Elektronik ve Bilgisayar Mühendisliği Sempozyumu, ELECO2010*, Bursa, Türkiye, 2010, pp. 515-515.

[2] American Cancer Society. Breast Cancer Facts&Figures 2011-2012 (11.07.2013).

[3] American Cancer Society. Cancer Facts&Figures 2012 (11.07.2013).

[4] V. Ozmen, "Meme hastalıklarının cerrahi tedavi kalitesi nasıl yükseltilebilir?", *The Journal of Breast Health*, vol. 5, no.2, pp. 119-121, 2009.

[5] *Türk Kanser Araştırma ve Savaş Kurumu Derneği, Türkiyede kanser istatistikleri*, http://www.kanser.gov.tr/index.php/daire-faaliyetleri/kanser-istatistikleri.html, (11.07.2013)

[6] S.J. Nass, I.C. Henderson, J.C. Lashof, "Mammography and Beyond: Developing Technologies for Early Detection of Breast Cancer", National Academy Pres, Washington, DC, 2001.

[7] J. Bozek, K. Delac, M. Grgic, "Computer-Aided Detection of Breast Abnormalities in Digital Mammography", 50th International Symposium ELMAR-2008, Zadar, Croatia, 2008, pp.45-52.

[8] A.L. Baert, M.F. Reiser, H. Hricak, M., Knauth, "Digital Mammography", Springer-Verlag Berlin Heidelberg, Berlin, Germany, 2010.

[9] S. Yu, L. Guan, "A CAD System for the Automatic Detection of Clustered Microcalcifications in Digitized Mammogram Films", *IEEE Transactions on Medical Imaging*, vol. 19, no. 2, pp. 115-126, February, 2000.

[10] B. Verma, J. Zakos, "A Computer-Aided Diagnosis System for Digital Mammograms Based on Fuzzy-Neural and Feature Extraction Techniques", *IEEE Transactions on Information Technology in Biomedicine*, vol. 5, no. 1, pp. 46-54, March, 2001.

[11] I. El-Naqa, Y. Yang, M.N. Wernick, N.P. Galatsanos, R.M. Nishikawa, "A Support Vector Machine Approach for Detection of Microcalcifications", *IEEE Transaction on Medical Imaging*, vol. 21, no. 12, pp. 1552-1563, December, 2002.

[12] L. Zhang, R. Sankar, W. Qian, "Advances in Micro-calcification Clusters Detection in Mammography", *Computers in Biology and Medicine*, vol. 32, no. 6, pp. 515-528, November, 2002.

[13] M. Masotti, "Exploring Ranklets Performances in Mammographic Mass Classification Using Recursive Feature Elimination", *Machine Learning for Signal Processing, Proceedings of the 2006 16th IEEE Signal Processing Society Workshop on*, 2006, pp.265-270.

[14] X. Zhang, X. Gao, Y. Wang, "MCs Detection with Combined Image Features and Twin Support Vector Machines", *Journal of Computers*, vol. 4, no. 3, pp. 215-221, March, 2009.

[15] MIAS Digital Mammogram database http://peipa.essex.ac.uk/info/mias.html

[16] S. Mukhopadhyay, B. Chanda, "Local Contrast Enhancement of Grayscale Images Using Multiscale Morphology", *Proc. ICVGIP-2000*, Bangalore, India, 2000, pp. 17-24.

[17] Z. Wei, Y. Hua, S. Hui-sheng, F. Hong-qi, "X-ray Image Enhancement Based on Multiscale Morphology", *Bioinformatics and Biomedical Engineering, ICBBE 2007, The 1st International Conference on*, Shanghai, China, 2007, pp. 702-705.

[18] R. C. Gonzalez, R.E. Woods, "Digital Image Processing", 2nd Edition, Prentice Hall, USA, 2002.

[19] S. Halkiotis, T. Botsis, M. Rangoussi, "Automatic Detection of Clustered Microcalcifications in Digital Mammograms Using Mathematical Morphology and Nerual Nerworks", *Signal Processing*, vol. 87, no. 7, pp. 1559-1568, July ,2007.

[20] R.M. Haralick, K. Shanmugam, I. Dinstein, "Textural Features for Image Classification", *IEEE Transactions on Systems , Man, and Cybernetics*, vol. Smc-3, no. 6, pp. 610-621, November, 1973.

[21] H.D. Cheng, M. Cui, "Mass Lesion Detection with a Fuzzy Neural Network", *Pattern Recognition*, vol. 37, no. 6, pp. 1189-1200, June, 2004.

[22] S.T. Bow, "Pattern Recognition and Image Preprocessing", Marcel Dekker, New York, USA, 2002.

[23] L. Costaridou, "Medical Image Analysis Methods", CRC Pres, USA, 2005.

[24] A. Meyer-Baese, "Pattern Recognition for Medical Imaging", Elsevier Academic Pres, California, USA, 2004.

[25] R.O. Duda, P.E. Hart, D.G. Stork, "Pattern Classification", Wiley-Interscience, New York, USA, 2001.

[26] J.A. Freeman, D.M. Skapura, "Neural Networks Algorithms, Applications, and Programming Techniques", Addison-Wesley Publishing Company , USA,1991.

[27] S. Haykin, "Neural Networks: A Comprehensive Foundation", Prentice Hall, New Jersey, USA,1999.