IMPROVING THE CLASSIFICATION ACCURACY OF COMPUTER AIDED DIAGNOSIS THROUGH MULTIMODALITY BREAST IMAGING

Nikolaos Pagonis[1], Dionisis Cavouras[2], Kostas Sidiropoulos[3], George Sakellaropoulos[1] and George Nikiforidis[1]

[1] Medical Image Processing and Analysis Group (MIPA), Laboratory of Medical Physics, School of Medicine, University of Patras, 26500 Rio, Greece

[2] Medical Image and Signal Processing Laboratory, Department of Medical Instrumentation Technology, Technological Educational Institute of Athens, 12210 Athens, Greece

[3] School of Engineering and Design, Brunel University West London, Uxbridge, Middlesex, UB8 3PH, UK, e-mail: Konstantinos.Sidiropoulos@brunel.ac.uk

e-mail: cavouras@teiath.gr, web page: http://medisp.bme.teiath.gr

Keywords: Image Analysis; Pattern Recognition; Multimodality; Breast Cancer; US X-RAY;

Abstract. *The purpose of the present study is to evaluate the effect of using multiple modalities on the accuracy achieved by a computer-aided diagnosis system, designed for the detection of breast cancer. Towards this aim, 41 cases of breast cancer were selected, 18 of which were diagnosed as malignant and 23 as benign by an experienced physician. Each case included images acquired by means of two imaging modalities: x-ray and ultrasound. Manual segmentation was performed on every image in order to delineate and extract the regions of interest (ROIs) containing the breast tumors. Then 104 textural features were extracted; 52 from the x-ray images and 52 from the US images. A classification system was designed using the extracted features for every case. Firstly, features extracted from x-ray images alone were used to evaluate the system. The same task was performed for features extracted from US images alone. Lastly the combination of the two feature sets, mentioned afore, was used to evaluate the system. The proposed system that employed the Probabilistic Neural Network (PNN) classifier scored 78.05% in classification accuracy using only features from x-ray. While classification accuracy increased at 82.95% using only features from US, a significant increase in the system's accuracy (95.12%) was achieved by using combined features from both x-ray and US. In order to minimize total training time, the proposed system adopted the Client-Server model to distribute processing tasks in a group of computers interconnected via a local area network. Depending on the number of clients employed, there was about a 4-fold reduction in training time employing 7 clients.*

1 INTRODUCTION

Breast cancer is the most common malignancy in women in the industrialized countries[1]. Early detection of breast cancer may lead to lower mortality rates. To achieve that, all palpable lesions must be examined. Additionally, every woman with suspicious inheritance must add precautionary inspection to her annual routine. Several imaging methods have been used to help in breast cancer diagnosis in early stages. Those methods include Digital X-ray mammography, ultrasound breast examination (US), Magnetic Resonance Imaging (MRI). Although MRI is considered the most sensitive method, it is not used frequently due to high cost. Instead, Digital mammography and US mammography are employed and are often used in a complementary manner, especially since ultrasound mammography involves non-ionizing radiation and it is considered as a reliable method [1]. Many studies have revealed the value of US in breast cancer diagnosis[1], while others claim that US is a valuable adjunct to X-ray. However, limited research has been done regarding multimodality breast imaging. Karen Drukker et al[2] implemented a computer-aided classification system combining features from mammography and ultrasound. Drukker proved that the accuracy of a CAD classification system can be improved by use of features from both modalities. The accuracy achieved by the system using both modalities was significantly higher than both accuracies achieved by the system using single modality data. Berkman Sahiner et al[3] designed a system called CADx that combines data acquired from US and X-ray mammography in order to improve the radiologists' performance in discriminating malignant from benign masses on mammograms and 3D ultrasound images.

The aim of the present study was to develop and evaluate a pattern recognition system, which uses textural features from two modalities (digital mammography and ultrasound), providing the physician with a convenient adjunctive tool capable to reduce false negative breast cancer detection. Features employed were extracted from the digital mammography and US images. In order to evaluate the efficiency of the proposed multimodality classification system, the following steps were performed. Firstly, features extracted from mammography images alone were used to evaluate the system. The same task was performed for features extracted from US images alone. Lastly the combination of the two feature sets was used to evaluate the system. The proposed system employed the Probabilistic Neural Network (*PNN*)[13] classifier. Classification performance was evaluated in terms of the derived accuracy.

2 MATERIAL AND METHODS

Forty one patients were examined by an experienced physician. From each patient US and X-Ray images were acquired. Tumor specimens from each patient were histollogically verified and 18 were diagnosed as malignant while 23 as benign. From each tumor on the digital images, a ROI was extracted, by means of specially designed software. From each ROI textural features were extracted and were subsequently used in the design of the classification system.

2.1 Feature extraction

Features comprised those calculated from the first order statistics (mean value, standard deviation, skewness, kurtosis)[18], the grey level spatial-dependence (GLSD) matrix[18] (angular second moment, contrast, correlation, inverse difference moment, entropy, sum average, sum variance, difference variance, difference entropy), the run-length (RL) matrix[19] (short run emphasis, long run emphasis, gray level non uniformity, run length non uniformity, run percentage) and the size and shape (area, perimeter, roundness, concavity) as shown in table I. Hence, each segmented ROI (15x15 pixel sub-image) was finally represented by a 52 feature vector (considering mean and range values for each feature from the co-occurrence and run length matrices). Thus, for each modality, two feature-classes were formed (malignant and benign), and a third dataset was additionally formed containing the features from both modalities.

All features were normalized to zero mean value and unit variance[17] according to $\tilde{x}_i = (x_i - m)/std$, where x_i and $\tilde{x}_i$ are the feature vectors prior and after the normalization, *m* and *std* are the mean value and standard deviation of each feature respectively, considering both classes.

A/A	Feature	A/A	Feature
1	Mean Value	14	Sum Variance
2	Standard deviation	15	Difference Entropy
3	Skewness	16	Difference Variance
4	Kurtosis	17	Absolute Value
5	Angular Second Moment	18	Short Run Emphasis
6	Contrast	19	Long Run Emphasis
7	Entropy	20	Grey Level non Uniformity
8	Inverse Difference Moment	21	Run Length non Uniformity
9	Autocorrelation	22	Run Percentage
10	Correlation	23	Area
11	Variance	24	Perimeter
12	Sum Average	25	Roundness
13	Sum Entropy	26	Concavity

Table 1: Extracted features

2.2 Classification

The PNN is a Bayes-Parzen classifier[23]. Equation 1 presents the discriminant function of a PNN classifier. Hence for class j:

$$g_j(x) = \frac{1}{(2\pi)^{p/2}\sigma^p N_j}\sum_{i=1}^{N_j} w(y_i) \qquad (1)$$

where $w(y_i)$ is a function of:

$$y_i = \frac{\sqrt{\|x - x_i\|^2}}{\sigma} \qquad (2)$$

In case

$$w(y) = e^{-\frac{y^2}{2}} \qquad (3)$$

which leads to a discriminant function with a Gaussian kernel:

$$g_j(x) = \frac{1}{(2\pi)^{p/2}\sigma^p N_j}\sum_{i=1}^{N_j} e^{-\frac{\|x-x_i\|^2}{2\sigma^2}} \tag{4}$$

where x denotes the test pattern vector to be classified, xi the i-th training pattern vector, Nj the number of patterns in class j, σ a smoothing parameter, and p is the dimensionality of the feature vector. The test pattern x is then classified under the class with the larger discriminant function value

2.3 Feature reduction and system evaluation

In order to reduce system dimensionality and computational time demands, features were reduced employing a statistical non-parametric test (wilcoxon) and the 10 most significant features (in terms of discriminatory power) were fed into the PNN classifier. Feature selection and individual classifier evaluation were performed by means of the exhaustive search algorithm[17] and the leave-one-out method[17], whereby all possible feature combinations were used to design and evaluate the classifier. The exhaustive search is preferred from other searching techniques (suboptimal) because it derives the most accurate results. The classifier was evaluated using features considering the digital mammography, the US, and both modalities, thus providing three overall accuracies. The best combination is that with the highest overall classification accuracy employing the least features.

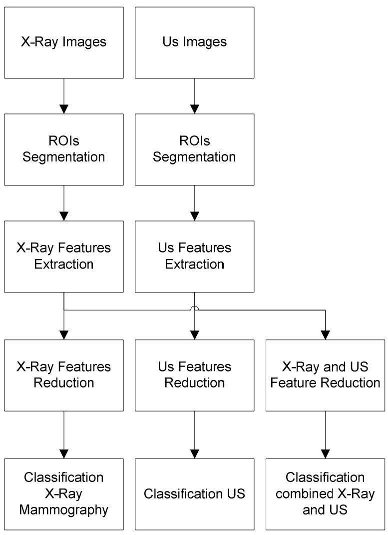

Figure 1: Flow chart of process

2.4. Parallel processing

In order to deal with the high computational burden introduced by the training and evaluation techniques involved, a parallel infrastructure was employed. By means of a custom made client-server application developed in JAVA programming language, the workload was efficiently distributed among a number of computers interconnected via a local area network, and total processing time was diminished. Specifically, each client was assigned with an even part of the large list of feature combinations derived by the exhaustive search method and reported the achieved accuracy back to the server.

3 RESULTS AND DISCUSSION

Classification accuracies of individual feature set are presented in Table 1. The classifier scored 78.05% overall accuracy using only features extracted from X-ray digitized images. Using features from US images the classifier scored 82.95%. The best result achieved from the X-ray and US features combination where the overall accuracy reached 95.12%.

Number of Features	**Accuracy (%)**		
	PNN		
	X-Ray	**US**	**Both modalities**
10	78.05	82.95	95.12

Table 1: Classification accuracies of individual feature sets.

From the present study some useful conclusions may be derived. The US breast imaging seems to be more sensitive than X-ray, a fact that has been already discussed by several researchers. The combination of multimodality data (X-ray and US) scored the highest accuracy, thus, suggesting that the best choice is to combine features from both modalities. Another issue is the computational demand of the proposed method. A reduction in the cost was accomplished by using computers interconnected into a local area network as illustrated in figure 2.

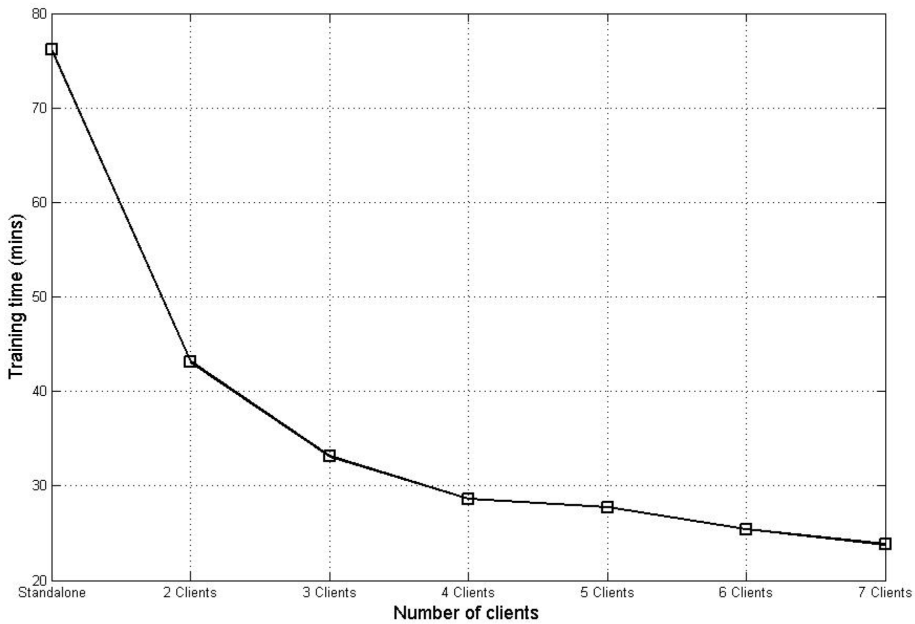

Figure 2: Number of clients versus training time.

4 CONCLUSION

The results of this study revealed that the use of features from both modalities from the same patient led to better accuracy results. It is common sense that the false negative mammogram result must be reduced or even vanished. Hence CAD systems designed to use more than one modality may be more objective as a second opinion diagnostic tool. Demanding training time may be reduced by distributing processing workload to a number of computers.

REFERENCES

[1] H.R. Laine, T. Tukeva , P. Mikkola , T. Holmstr6m B.W., Kleihues, P. (1995), "Assessment Clinical paper of mammography and ultrasound examination in the diagnosis of breast cancer", Elsevier.

[2] Karen Drukker, PhD, Karla Horsch, PhD, Maryellen L. Giger, PhD (2005), "Multimodality Computerized Diagnosis of breast lesions using mammography and sonography", Acad Radiol 2005, 12:970–979.

[3] Berkman Sahiner PhD, Heang-Ping Chan PhD, Lubomir M. Hadjiiski PhD, Marilyn A. Roubidoux MD, (2008) "ROC Study of the Effect on Radiologists' Accuracy in Characterizing Breast Masses on Mammograms and 3D Ultrasound Images" Academic Radiology 2009:1–9.

[4] Ioannis Pratikakis, Anna Karahaliou, Katerina Vassiou, Vassilis Virvilis, Dimitrios Kosmopoulos, and Stavros Perantonis (2007), "eMedI:Web-Based E-Training for Multimodal Breast Imaging", Proceedings of world academy of science, engineering and technology volume 25 november 2007 issn 1307-6884

[5] Diaz, L.K., Sneige, N. (2005), "Estrogen receptor analysis for breast cancer: Current issues and keys to increasing testing accuracy", Adv Anat Pathol, Vol. 12, pp. 10-19.

[6] Diaz, L.K., Sahin, A., Sneige, N. (2004), "Interobserver agreement for estrogen receptor immunohistochemical analysis in breast cancer: A comparison of manual and computer-assisted scoring methods", Ann Diagn Pathol, Vol. 8, pp. 23-27.

[7] Mofidi, R., Walsh, R., Ridgway, P.F., Crotty, T., McDermott, E.W., Keaveny, T.V., Duffy, M.J., Hill, A.D., O'Higgins, N. (2003), "Objective measurement of breast cancer oestrogen receptor status through digital image analysis", Eur J Surg Oncol, Vol. 29, pp. 20-24.

[8] Schnorrenberg, F., Tsapatsoulis, N., Pattichis, C.S., Schizas, C.N., Kollias, S., Vassiliou, M., Adamou, A., Kyriacou, K. (2000), "Improved detection of breast cancer nuclei using modular neural networks", IEEE Eng Med Biol Mag, Vol. 19, pp. 48-63.

[9] Lehr, H.A., Mankoff, D.A., Corwin, D., Santeusanio, G., Gown, A.M. (1997), "Application of photoshop-based image analysis to quantification of hormone receptor expression in breast cancer", J Histochem Cytochem, Vol. 45, pp. 1559-1565.

[10] Bejar, J., Sabo, E., Misselevich, I., Eldar, S., Boss, J.H. (1998), "Comparative study of computer-assisted image analysis and light-microscopically determined estrogen receptor status of breast carcinomas", Arch Pathol Lab Med, Vol. 122,

pp. 346-352.

[11] Makkink-Nombrado, S.V., Baak, J.P., Schuurmans, L., Theeuwes, J.W., van der Aa, T. (1995), "Quantitative immunohistochemistry using the cas 200/486 image analysis system in invasive breast carcinoma: A reproducibility study", Anal Cell Pathol, Vol. 8, pp. 227-245.

[12] Gose, E., Johnsonbaugh, R., Jost, S. (1996), *Pattern recognition and image analysis*. Prentice Hall PTR, New Jersey.

[13] Specht, D.F. (1990), "Probabilistic neural networks", Neural Networks, Vol. 3, pp. 109-118.

[14] Gonzalez, R., Woods, R. (2002), *Digital image processing*. Addison-Wesley Pub, NY.

[15] Kostopoulos, S., Cavouras, D., Daskalakis, A., Ravazoula, P., Nikiforidis, G. (2006), "Image analysis system for assessing the estrogen receptor's positive status in breast tissue carcinomas", *Proceedings of International Special Topic Conference on Information Technology in Biomedicine, Ioannina Greece,*

[16] Glotsos, D., Spyridonos, P., Petalas, P., Cavouras, D., Ravazoula, P., Dadioti, P.A., Lekka, I., Nikiforidis, G. (2004), "Computer-based malignancy grading of astrocytomas employing a support vector machine classifier, the who grading system and the regular hematoxylin-eosin diagnostic staining procedure", Anal Quant Cytol Histol, Vol. 26, pp. 77-83.

[17] Theodoridis, S., Koutroumbas, K. (2003), *Pattern recognition*. Elsevier, San Diego.

[18] Haralick, R.M., Shanmugam, K., Dinstein, I.H. (1973), "Textural features for image classification", IEEE Trans Sys Man Cyb, Vol. 3, pp. 610-621.

[19] Galloway, M.M. (1975), "Texture analysis using gray-level run lengths", Computer Graphics and Image Processing, Vol. 4, pp. 172-179.

[20] Chappard, D., Pascaretti-Grizon, F., Gallois, Y., Mercier, P., Basle, M.F., Audran, M. (2006), "Medullar fat influences texture analysis of trabecular microarchitecture on x-ray radiographs", Eur J Radiol, Vol. 58, pp. 404-410.

[21] Looney, S.W. (1988), "A statistical technique for comparing the accuracies of several classifiers", pattern Recognition Letters, Vol. 8, pp. 5-9.

[22]M. Hajmeer, I. Basheer, A probabilistic neural network approach for modeling and classification of bacterial growth/no-growth data, Journal of Microbiological Methods 51, 217– 226, (2002)

[23] Masters T., Advanced Algorithms for Neural Networks Wiley, New York, 1995.