thermal image processing,
neural network, classification

T. JAKUBOWSKA[*], B. WIECEK[**], M. WYSOCKI[**],
C. DREWS-PESZYNSKI [*], M. STRZELECKI[**]

CLASSIFICATION OF BREAST THERMAL IMAGES USING ARTIFICIAL NEURAL NETWORKS

In this paper we present classification of the thermal images in order to discriminate healthy and pathological cases during breast cancer screening. Different image features and approaches for data reduction and classification have been used to distinguish healthy breast one with malignant tumour. We use image histogram and co-occurrence matrix to get thermal signatures and analyze symmetry between left and right side. The most promised method was based on wavelet transformation and nonlinear neural network classifier. The proposed approach was used in the pilot investigations in the medical centre which is permanently using thermograph for breast cancer screening, as an adjacent method for other classical diagnostic method, such as mammography.

1. INTRODUCTION

Breast cancer and other breast diseases are a major issue in women's health today, particularly in advanced industrialized nations. It takes years for a tumour to grow, and the earliest possible indication of abnormality is needed to allow for the earliest possible treatment and intervention. Thermal imaging can be useful as early diagnostic technique that can detect breast cancer.

At the Technical University of Lodz, in Institute of Electronics, the software for calculation image features, their selection and processing, as well as classification algorithm has been developed [2, 3, 6, 10]. Among the variety of different image features, statistical thermal signatures:

- histogram-based (mean, variance, skew ness, kurtosis and five histogram percentiles);
- co-occurrence matrix based features which are defined for matrices constructed for five distances between image pixels (d=1, 2, 3, 45), and separately for the four directions: horizontal, vertical, 45 and 135 (angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy);

[*] Laser Diagnosis and Therapy Centre at Technical University of Lodz, Poland
[**] Institute of Electronics, Technical University of Lodz, Poland

- gradient-based features (absolute gradient mean, variance, skew ness, kurtosis, and of non-zero gradients);
- run-length matrix-based features (short run emphasis inverse moment, long run moment, grey level no uniformity, run length no uniformity and fraction of in runs, for the same directions as in the case of co-occurrence matrix based features).

The next group if methods are based on image transformations, such as linear filtering, Fourier or wavelet analysis. All these methods allow to regenerate an image which is processed or converted, and signatures are defined in different domain, eg.: space frequency or scale domain. In the research presented here some new features based on wavelet transformation were introduced.

It is possible to define hundreds of features for an image, and obviously, the selection and reduction are needed. Two approaches were applied in this work, based on Fischer coefficient and by using minimization of classification error probability (POE) and average correlation coefficients (ACC) between chosen features [3]. It can reduce the number of features to a few ones, maximum to 10. The next step is the features pre-processing which generates new parameters after linear or nonlinear transformations. It allows to get data which less correlated and of the lower order. Two approaches were used in the research, i.e.: PCA (Principal Component Analysis) and LDA (Linear Discriminate Analysis) [3, 7].

Finally, classification was performed using different Artificial Neural Network (ANN), with or without additional hidden layers, and with different number of neurons. Additionally, we have compared ANN classification with widely used Nearest Neighbour Classification (NNC).

2. IMAGE PROCESSING

Wavelet transformation is actually used in many domains, such as telecommunication and signal processing for compression and to extract quantitative data from a signal. In image processing it can be employed to get new features, representing both global and detail information. Wavelet transformation is based on image filtering represented by rows and columns using low and high pass linear filters (Fig.1). After filtering, decimation is used to reduce number of pixels. The procedure can be repeated until 1x1 images are obtained. Practically, the processing is stopped earlier, after 2-4 steps, and then the features are derived from the filtered subimages.

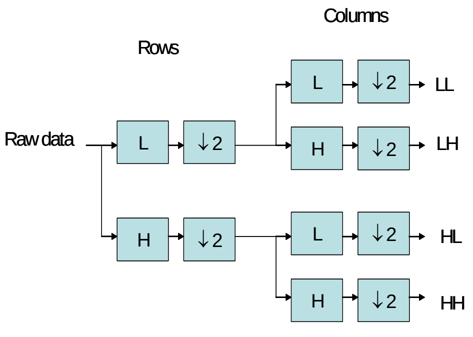

Fig. 1. Wavelet transformation of an image

Artificial neural network is used for classification. The selected image features has been used as inputs. It means that the number of inputs nodes is equal to number of the features. Number of neuron in the first hidden layer can be equal or lower than the number of features in the classification, as shown in Fig.2. ANN can have user-defined next hidden layers which allow additional nonlinear processing of the input features. As ANN is the nonlinear system, such technique allows to additional decorrelating and data reduction, what finally improves the classification. Such approach is known as Nonlinear Discriminant Analysis (NDA) [11-12].

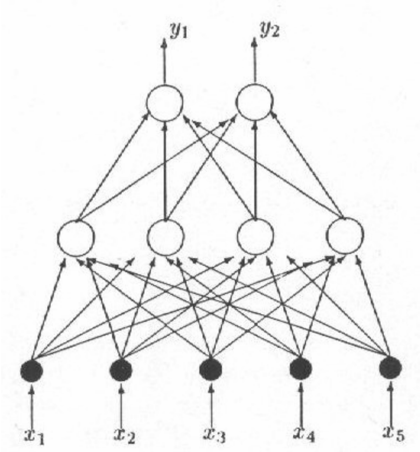

Fig. 2. Neural network example with input, single hidden and output layers

It is well known that the training of ANN is the very important step in the entire protocol. It is an multivariable optimization problem based on typically back propagation technique. In general case it can lead to wrong solutions if the there is no single minimum of the error function. That is why we need enough data during learning phase, and sometimes it is better or necessary to repeat training of ANN with different initial values of the neuron weight coefficients.

3. RESULTS

In the presented investigations, 10 patients with recognized breast tumors, as well as 30 healthy patients were used. All healthy and pathological cases were confirmed by other diagnostic methods, such as mammography, USG, biopsy, etc. We have used thermographic camera to take 2 images for each breast: frontal and side ones. Each patient was relaxing before the experiment for 15 min. in a room where temperature were stabilized at 20^{o}C.

Then the numerical procedure was used to calculate features of the images. Fig.3 presents example of thermal image of healthy breast. Fig.4 and Fig.5. present the pathological cases where breast with tumor has evidently higher temperature and the temperature distribution is very asymmetric.

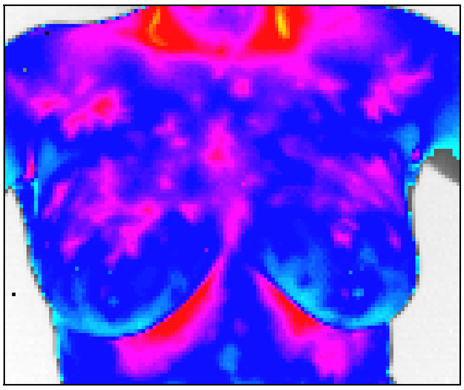

Fig. 3. Example of thermal image of healthy breast

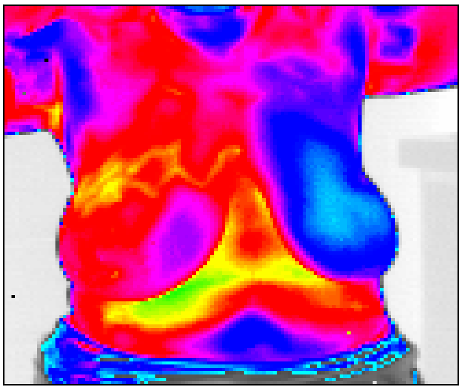

Fig. 4. Thermal image of the breast with malignant tumor (yellow area on the left side)

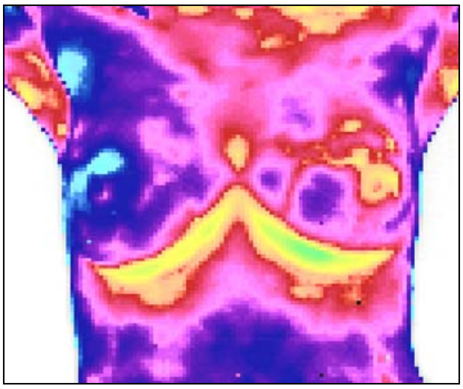

Fig. 5. Thermal image of the breast with malignant tumor (orange area on the right side)

One of the possible transformations uses wavelets realized by filtering as it has been mentioned above. The result of such processing is presented in Fig.6. As it is seen, the different filters are showing different details of the image, i.e.: high pass filters allows to present gradient of the temperature, while the low pass one displays the global temperature distribution and the energy of the signal understood as the level of temperature.

As a powerful tool, wavelet transformation gives the possibility of generating features from different subimages, in different scales and subbands. Filtering can be easily parameterized in the sense of varying cut-off frequency, what provides the additional flexibility in the algorithm. It denotes that hundreds of different features can be produced by this method. Obviously, the selection of these features is necessary to obtain only these ones which are the most discriminative and weakly correlated [6].

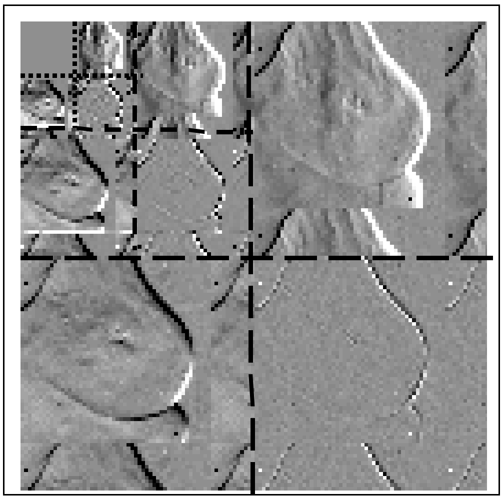

Fig. 6. Result of wavelet transformation

Classification can start from the raw data analysis. As seen in Fig.7, we have two classes containing non-healthy (1) and healthy (2) cases. Selection of features reduced its number to three – two first derived from so-called co-occurrence matrix, (sum of squares and inverse difference moment) [4, 10]. The third one is based on the wavelet transformation – energy E_{HL} for the scale no.1 and high and low frequency subbands. Raw data analysis does not give satisfactory feature separation and classification. It is hard to separate clusters of features corresponding to physiological and pathological images. The distance in between them in the multivariate space is not large enough, what means that the probability of erroneous classification is quite high.

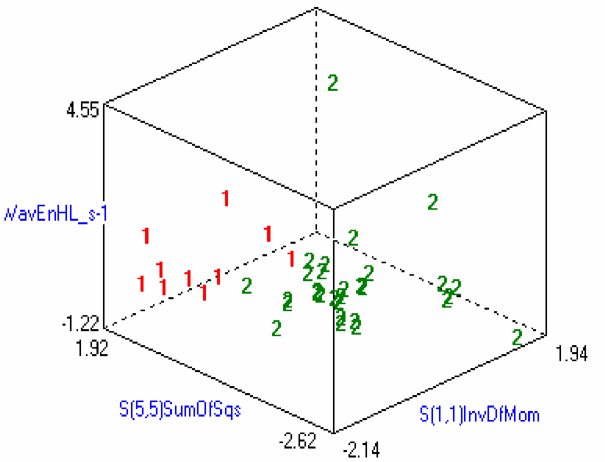

Fig. 7. Raw data analysis result

The next results were obtained by LDA (Fig. 8), in which the most discriminative features (MDF) are generated [3]. LDA creates typically new set with smaller number of features. LDA is the linear transformation, and produces linearly separated features, what means that in general case it is also no possible to separate them fully. This case is illustrated in Fig. 8.

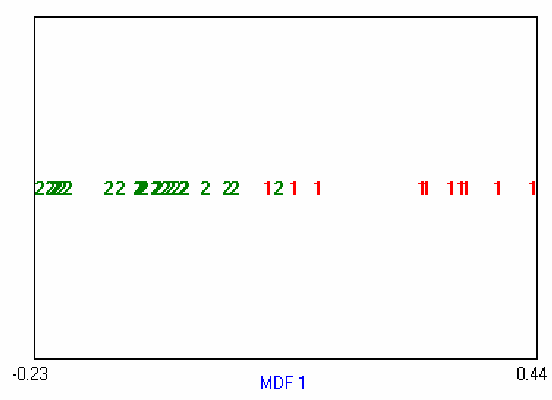

Fig. 8. LDA non-acceptable separation of cluster

Nonlinear transformation of the features obtained by ANN with additional hidden layer is the most promising technique. The output of this additional hidden layer creates the new smaller set of features which are typically much better separable. It can be simple verified by the larger value of Fisher coefficient. In our investigation 7 original features were selected using POE and ACC selection criteria. First and second hidden layers contained only one and two neurons, respectively. It denotes that the original features were reduced to 2 new ones. Our expectation that they allow better separation and classification was confirmed.

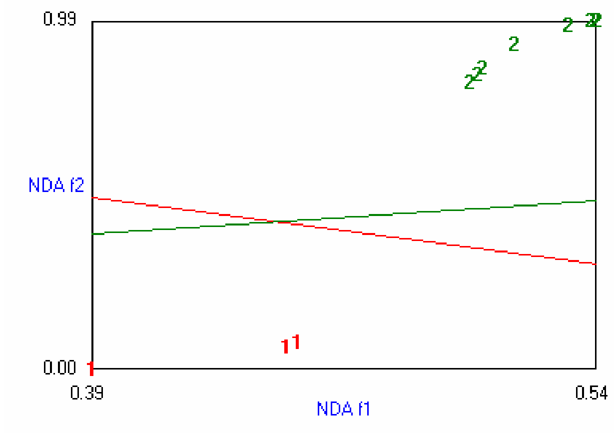

Fig. 9. NDA classification results for side thermal images of the breast

The results of NDA are presented in Fig. 9-10. Two new features are far away from each other on the feature space, both for frontal and side images. It was not very surprising that value of Fisher coefficient F=2.7 for original features was increased to F = 443.4. ANN together with NDA have one more advantage. Besides of data reduction and decorrelating, it also allows implementing classification realized by the last output layer of ANN.

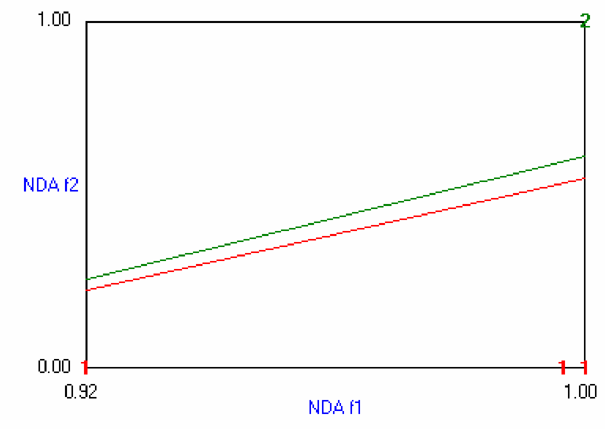

Fig. 10. NDA classification results for frontal thermal images of the breast

We faced one problem in the presented research. The number of pathological cases was small, only 10. As it has been already mentioned it is difficult to get well diagnosed patients, especially in the young age [1, 9]. The research is still in progress, and we actually enlarging our image database day-by-day. This research is a preliminary one, and shows the possibility of using advanced image processing tools. The effectiveness the screening procedure can be verified later, when we will have enough input data.

Because of these limits above, the classification was carried out in the following the simplified way. Every image from each class (healthy and non-healthy) was extracted, and the training of ANN was performed for all the rest. Then this one not being used before during learning was employed now for classification.

The results are presented in Table I. First 3 rows in this Table present Nearest Neighbour (NN) classification of the image features. The first results were obtained by using feature calculated on an image and NN classification after feature selection. In the next experiments, pre-processing of original features using PCA and LDA was employed.

Table 1. Errors of classification

	Frontal positions		Side positions	
	False negative	False positive	False negative	False positive
Raw data	2/10	3/30	1/10	2/30
PCA	2/10	3/30	1/10	2/30
LDA	1/10	2/30	1/10	1/30
NDA	2/10	2/30	0/10	2/30

Finally we performed calculation using NDA. Additional hidden layer consisted of 2 neurons for generating 2 new more discriminative features. The results in Table I confirm the effectiveness of using both NN and ANN classifiers. Although it was not evidently proved that NDA results are better in the classification, we definitely conclude that ANN is a powerful tool for thermal image processing during breast cancer screening. Reducing false/positive errors of classification is the most important task for the future research.

4. CONCLUSION

This paper presents the results of the feature analysis for thermal images used in breast oncology. Thermography as the additional and adjacent method can be very helpful for early and image classification that helps to recognize tumors. Although we do not have many pathological cases for investigations yet, the first results are very promising.

Breast cancer screening is a challenge today for medical engineering. Breast temperature depends not only because of some pathological changes, but it also varies in normal physiological situations, even it is a consequence of emotional state of a patient (Fig.11). It was a main reason that we were looking for an advanced method of image processing, that could give satisfactory results.

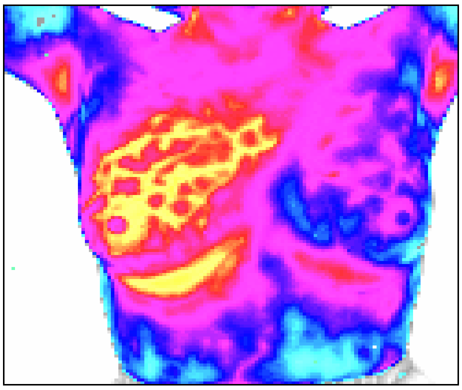

Fig. 11. Thermal image of an patient during lactation.

One of the possible alternative for such processing in Artificial Neural Network Classification based on multidimensional feature domain, with use of modern transformations, e.g. on wavelets. The preliminary investigations are quite successful, and can be improved by increasing the number of samples taken for processing.

Future research will concentrate around selection of features and adjusting wavelet transformation parameters to get the best classification. We assume that the more satisfactory results can be obtained by using features based on asymmetry between left and right side of a patient. It could help for one-side cancerous lesion classification, what is the most typical pathological case and frequently happens today.

BIBLIOGRAPHY

[1] BENNETT M. “Breast Cancer Screening Using High-Resolution Digital Thermography”, Total Health, Vol. 22 No 6 p.44, 1985.

[2] CAUSTON D. R., “A Biologist's Advanced Mathematics”, London, Allen and Unwin, 1987.

[3] DEBIEC P., STRZELECKI M., MATERKA A., "Evaluation of Texture Generation Methods Based on CNN and GMRF Image Texture Models", International Conference on Signals and Electronic Systems ICSES'2000, Ustron, Oct. 2000, pp. 187-192.

[4] JAKUBOWSKA T., WIECEK B., WYSOCKI M., DREWS-PESZYNSKI C., “Thermal Signatures for Breast Cancer Screening Comparative Study, Proc. IEEE EMBS Conf. Cancun, Mexico, Sep 17-21, 2003.

[5] JOLLIFFE I. T., “Principal Component Analysis”. New York, Springer-Verlag, 1986.

[6] KOCIOLEK M., MATERKA A., STRZELECKI M., SZCZYPINSKI P., "Discrete Wavelet Transform-Derived Features for Digital Image Texture Analysis", Proc. International Conference on Signals & Electronic Systems ICSES'2001, Lodz, 18-21 September 2001, pp. 111-116

[7] MANLY B. F. J, “Multivariate Statistical Method: A Primer”. Chapman & Hall, London, 1994.

[8] MATERKA A., STRZELECKI M., LERSKI R., SCHAD L., „Evaluation of Texture Features of Test Objects for Magnetic Resonance Imaging”, Infotech Oulu Workshop on Texture Analysis in Machine Vision, June , 1999, Oulu, Finland.

[9] NG E.Y.K., UNG L.N., NG F.C., SIM L.S.J. “Statistical Analysis Of Healthy And Malignant Breast Thermography”, Journal of Medical Engineering & Technology, Vol. 25 No 6 (Nov/Dec 2001) p.253-263.

[10] SCHÜRMAN J. (1996) Pattern classification, John Wiley & Sons, 1996