Neural Network-based Classification of Signal-Time Curves Obtained from Dynamic MR Mammographic Image Series

R. Lucht [1], M.V. Knopp [2], G. Brix [1]
[1] Institute of Radiation Hygiene, Federal Office for Radiation Protection, Neuherberg, Germany and
[2] 'Radiological Diagnostics and Therapy', German Cancer Research Center, Heidelberg, Germany

Introduction

Dynamic contrast-enhanced magnetic resonance imaging (dMRI) has been proven to increase sensitivity and specificity in the detection of breast cancer. Two analysis methods for the signal-time courses were suggested: (a) definition of threshold values and (b) pharmacokinetic evaluation. While the first method did not use all available information, the second requires a-priori knowledge on tissue microcirculation. A neural network approach by-passes these restrictions. By presenting examples of known class specification, the neural network is able to learn the distribution patterns of „typical" features of each class (supervised learning). Afterwards, the network is able to separate non-determined data into the different classes and to calculate the probability of the data to belong to each of the classes. The goal of our study was to evaluate the role of neural networks in the classification of signal-time curves obtained from dMRI data sets of the female breast. Special emphasis was placed on the question of whether a low temporal resolution would result in sufficient values for sensitivity and specificity.

Materials & Methods

All patient examinations were performed at a 1.5-Tesla whole-body MR system (Magnetom SP). The patient lies prone in a double-breast coil with the arms extended above the head. Saturation-Recovery-TurboFLASH images (recovery time T_{REC}= 125 ms, TR= 9 ms, TE= 4 ms, α= 12°, raw matrix 128×128, image matrix 256×256, TH= 6 mm, FOV= 320 mm, 15 slices) were acquired with a repetition rate of 23 s over a period of 12 min, resulting in a dynamic image series of 32 data sets. The contrast agent Gd-DTPA (Magnevist; dosage 0.1 mmol/kg body weight) was applied at the beginning of the fifth repetition cycle with a constant infusion rate over 1 min.

The signal-time courses in each voxel of the dMRI data sets were evaluated by a pharmacokinetic approach [1,2,3] to identify suspect lesions. Within these suspicious areas, regions of interest (ROIs) were defined yielding signal-time curves. All examined lesions were histopathologically classified (162 malignant, 102 benign) as being carcinoma, fibroadenoma or mastopathia. Additionally, 105 parenchyma signal-time curves were acquired. Thus, the signal-time curves belong to four tissue classes: „carcinoma", „fibroadenoma", „mastopathia", and „parenchyma". In a second attempt, the classes „fibroadenoma" and „mastopathia" were joined to form a new class, called „benign". 70% of the overall number of data sets were used for training of the network and 30% for testing the final performance.

The signal-time curves were normalized to the average of the pre-contrast intensities. The pre-contrast measurements were excluded from the analysis, resulting in signal-time curves with 28 measurement points. To evaluate different sampling strategies, the number of data points were reduced to (a) 14 measurement points, taking every second measurement point, (b) 7 measurement points, one every minute, and (c) 3 measurement points measured 1.75, 3.2, and 10 minutes after the beginning of contrast medium administration. The following eight multi-layer perceptrons (MLP) were evaluated: MLP-28-4-4, MLP-14-4-4, MLP-7-4-4, MLP-3-4-4, MLP-28-4-3, MLP-14-4-3, MLP-7-4-3, MLP-3-3-3, whereby the three numbers for each network define the number of input, hidden, and output nodes. Each of the MLPs was trained 20 times by presenting the curves belonging to the training data set in different order. This procedure yielded a receiver operation curve (ROC) for each network. For comparison, the signal-time curves were classified by an experienced radiologist.

Results and Discussion

More than 87% of parenchyma signal-time curves were correctly classified, even by the network classifier with only three input nodes. The values for sensitivity and specificity in discrimination between malignant and benign lesions (carcinoma versus fibroadenoma plus mastopathia) are best in the case of the network MLP-28-4-3 (sensitivity 83.7%, specificity 81.4%) and decrease slightly with smaller numbers of input nodes. But MLP-3-4-4 still has a performance level of 76% for sensitivity and 76% for specificity. For all MLPs, the ROC for the networks with four output nodes (carcinoma, parenchyma, fibroadenoma, mastopathia) are lower than the ones obtained by the networks with three output nodes (carcinoma, parenchyma, benign). Poor results were achieved in differentiation between fibroadenomas and carcinomas plus mastopathia, and mastopathia and carcinoma plus fibroadenoma. Most misclassified fibroadenomas were considered to be mastopathias.

The characterization of the signal-time curves of our test data set as most likely being malignant by an experienced radiologist in a blinded reader analysis yielded a sensitivity of 85.4%, specificity of 81.8% and overall accuracy of 84.3%. Two recent studies, in which the discrimination between signal-time curves from malignant and benign lesions was carried out on the basis of visual inspection of the curves without giving consideration to morphological or pharmacokinetic criteria, yielded similar results. Knopp et al. [4] obtained values of 69% for sensitivity and 93% for specificity, whereas Daniel et al. [5] obtained values of 82% for sensitivity and 62% for specificity. These values correspond to the ROC obtained by the neural network analysis.

In conclusion, neural networks have the potential to identify malignant lesions with nearly the same accuracy as experienced human readers even if the number of measurements is reduced. Moreover, they are able to quantify the reliability of the classification result in terms of probabilities. A more precise specification of the type of a particular benign lesion, can barely be obtained even with 28 measurement points, without using additional morphological or pharmacokinetic information. A pixel-by-pixel classification (segmentation) of dynamic mammographic images with neural networks can be used to highlight areas with suspicious signal intensity courses.

References

1. Knopp, M.V., et al., *MRI Clinics of North America,* 2, 633, 1994
2. Hoffmann, U., et al., *Magn. Reson. Med.,* 33, 506, 1995
3. Brix, G., et al., *J. Comput. Assist. Tomogr.*,15, 621, 1991
4. Knopp, M.V., et al., *Radiologe, 38, 307,* 1998
5. Daniel, B.L., et al., *Radiology* 209, 499, 1998