ELSEVIER
Artificial Intelligence in Medicine xxx (2003) xxx–xxx
Artificial Intelligence in Medicine
www.elsevier.com/locate/artmed

WeAidU—a decision support system for myocardial perfusion images using artificial neural networks

Mattias Ohlsson*

Department of Theoretical Physics, Lund University, Sölvegatan 14A, SE-223 62, Lund, Sweden

Received 3 January 2002; received in revised form 1 April 2003; accepted 14 April 2003

Abstract

This paper presents a computer-based decision support system for automated interpretation of diagnostic heart images (called WeAidU), which is made available via the Internet. The system is based on image processing techniques, artificial neural networks (ANNs) and large well-validated medical databases. We present results using artificial neural networks, and compare with two other classification methods, on a retrospective data set containing 1320 images from the clinical routine. The performance of the artificial neural networks detecting infarction and ischemia in different parts of the heart, measured as areas under the receiver operating characteristic curves, is in the range 0.83–0.96. These results indicate a high potential for the tool as a clinical decision support system.

Keywords: Artificial neural networks; Diagnosis; computer-assisted; Myocardial infarction; Myocardial ischemia; Myocardial perfusion images

1. Introduction

The practice of medicine is to a great extent an information-management task. The decision making by a physician is based upon expert knowledge, information from the individual patient and information from many previous patients, the latter known as experience. The interpretation of patient data is difficult and complicated. Mainly, because the required expert knowledge in each of many different medical fields is enormous and the information available for the individual patient is multi-disciplinary, imprecise and very often incomplete. The interpretation of the available data from a patient is made by a physician, who may have a limited knowledge, and experience in analysis of only some of

* Tel.: +46-222-7782; fax: +46-222-9686.
E-mail address: mattias@thep.lu.se (M. Ohlsson).

0933-3657/$ – see front matter © 2003 Published by Elsevier Science B.V.
doi:10.1016/S0933-3657(03)00050-2

the data. In this situation physicians can usually consult a more experienced colleague at the clinic.

With the aid of modern information technology, it is possible for a physician to contact an experienced colleague at another hospital and transfer data to him or her. For example, a physician at a remote hospital can send a diagnostic image to an experienced physician at a university hospital. Thereafter they can discuss the image over the phone. A problem with this technique is that experienced physicians are not always available when the advice is needed. Therefore computer-based decision support systems available via the Internet is an interesting alternative. With this technique decision support is available 24 h per day, 365 days per year.

WeAidU, the project described in this paper, is a computer-based decision support system for automated interpretation of myocardial perfusion images. Such heart images can be used to diagnose myocardial infarction or ischemia. The system uses artificial neural networks (ANNs) as classification tool.

The use of ANNs to interpret myocardial perfusion images has appeared in several studies [3,4,10,16]. The best of these ANNs has shown to perform as well as or even better than human experts. Despite their high performance, this type of ANNs will not take over the decision-making process from the physician in clinical practice. Rather, the computer can assist the physician by proposing an interpretation of the scintigram. A recent study by Lindahl et al. [11] showed that physicians benefit from the advice of ANNs both in terms of an improved performance and a decreased intra- and inter-observer variability. It has also been shown that this type of ANNs can maintain a high accuracy in a hospital separate from that in which they were developed [12].

The outline of this paper is as follows: Section 2 contains a description of the heart images that the WeAidU system analyzes. These images are preprocessed by a Fourier method as described in Section 2.2, followed by a discussion about the gold standard used. A description of the ANN classifiers used in the WeAidU system is found in Section 2.5. The performance of the ANN classifiers, and two other classification methods, are presented in Section 3 together with a short technical description of the WeAidU system, including an image of the interface. The paper is ended with a discussion.

2. Methods

2.1. Myocardial perfusion images

The blood-flow to the heart can be examined by injecting a radioactive tracer (e.g. ^{99m}Tc-sestamibi) and thereafter acquiring scintigraphic images with a gamma camera. The patient is usually examined both at rest and after exercise and the result can be presented as two so called bulls-eye images, see Fig. 1. Different regions in the bulls-eye image corresponds to different regions of the heart, where the upper part, for instance, corresponds to the anterior wall of the heart.

Dark areas in these images represent parts of the heart with reduced blood-flow, generally caused by coronary artery disease (CAD). Differences between the rest and the stress images may indicate that the patient suffers from ischemia. The interpretation of

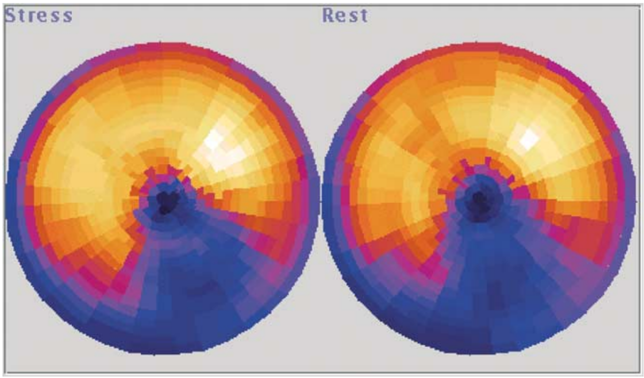

Fig. 1. Heart images in the form of so called bulls-eye images, one obtained at rest (right figure) and the other after exercise (left figure).

these images is a pattern recognition task. The physician must rely on his or her experience rather than on simple rules of how to interpret the images. A less experienced physician can benefit from a computer-based decision support system when a more experienced colleague is not present. Also a very experienced physician can use such a system for a second opinion.

2.2. Image preprocessing

The size of the bulls-eye images are 17×64 pixels. Image preprocessing is therefore used to decrease the amount of data for each image and to extract relevant features. A Fourier transform technique was used as follows: each of the two 17×64 images (rest and stress) was expanded by mirroring about row 17, and then discarding the last row (i.e. the first row of the succeeding Fourier period), to produce 32×64 matrices. The two 32×64 matrices were input as the real and imaginary parts of a complex 32×64 matrix in a fast Fourier transform [17]. The spatial low-frequency components describing the rest and stress images are found near the origin in the spatial frequency plane, as mapped by the transformed complex matrix. Selections of up to 36 complex low-frequency components (72 real values) were made for further image analysis. Since the relative importance of axial and radial information is a priori unknown, different sets of components were tried. A selection of 30 values constituting the real and imaginary part of the coefficients for 15 of the lowest frequencies were used as ‘‘feature vectors’’ for each pair of rest and stress images. For more details see [10].

2.3. Study population

The first studies by Lindahl et al. [10,11] used a study population of 135 images recorded at Lund University Hospital during the period from November 1992 to October 1994. During the course of development more images have been collected, both from hospitals in Sweden and other European countries. Currently the database consists of about 1500 myocardial perfusion images where many different patient examination modalities are represented, such as different types of radioactive tracers used in the myocardial scintigraphy. Three different tracers are represented in the database, namely, ^{99m}Tc-sestamibi, ^{99m}Tc-tetrofosmin and ^{201}Th. There are many different cameras that can be used when acquiring the scintigraphy data and the database (currently) covers four different such cameras (General Electric, Toshiba, Picker and Siemens). In the *one-day protocol* the rest and stress-examinations are performed on the same day, whereas in the *two-day protocol* they are split into two days. Both protocols are present in the database.

We believe that this database represents a proper selection of the perfusion images found in the daily clinical practice. Thus, a decision support system based on this database can therefore be expected to generalize to multiple clinical centers. More images will, however, be collected to further enlarge the database.

The results, presented in this paper, from the comparison between different classification methods are based on a subset of 1320 patients coming from four different hospitals in Europe.

2.4. Gold standard

In order to train the ANN classifiers a gold standard is needed. In this paper, we used a gold standard taken from the clinical routine evaluation of the patients. The physicians that performed these evaluations are very experienced and can be regarded as experts in their field.

Coronary angiography is a more objective gold standard and is commonly used in studies concerning myocardial perfusion images. One disadvantage with coronary angiography is that the angiogram shows different phenomenon compared to perfusion images. For instance, a stenosis in a coronary artery does not always correlate with a reduction in myocardial perfusion. The clinical evaluation has the advantage of being more close to the clinical reality. One disadvantage however, is the variability among different physicians and the variability for the same physician, which makes the classification problems more difficult.

2.5. Artificial neural networks

The WeAidU decision support system currently delivers two diagnostic advice, one regarding the presence of infarction and one that concerns ischemia. Furthermore, the heart is divided into five physical territories (anterior, inferior, apical, septal and lateral) and a diagnostic advice is given for each of these territories (see Fig. 2). The system uses 10 different ANN classifiers one for each of the advice given. Each classifier consists of an ensemble of single ANNs (see Section 2.5.2).

The individual members of the ensemble are standard multi-layer perceptrons [19] with one hidden layer consisting of 5–15 nodes and one output node that encodes infarction/ischemia or not. The tanh() activation function is used in the hidden layer and for the output layer the logistic function is used. Each multi-layer perceptron is trained using gradient descent applied to a cross-entropy error function. The gradient descent method is

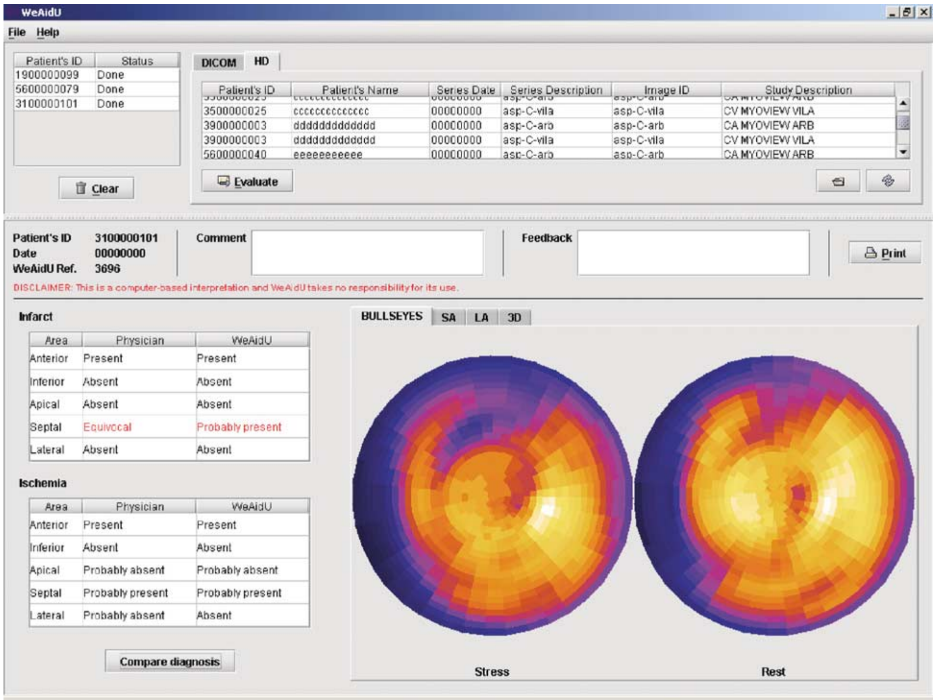

Fig. 2. The WeAidU interface. The upper part shows the selection of different studies, either using files stored on the local hard disk or using the DICOM network protocol. The lower part shows the interpretation of the WeAidU system (WeAidU column) and the mandatory interpretation suggested by the physician (physician column), for each of the five territories of the heart concerning infarction and ischemia.

augmented with a traditional momentum term and a Langevin extension [20]. The Langevin extension consists of adding a random Gaussian component to the weight updates, which has the effect of speeding up the minimization procedure.

2.5.1. Neural network regularization

In order to avoid over-training a weight elimination [6] regularization term of the following form is used,

$$\alpha \sum_i \frac{w_i^2}{1 + w_i^2} \tag{1}$$

where the sum runs over all weights w_i in the multi-layer perceptron, excluding the threshold weights since they should not be part of the regularization. This term penalizes large weights, resulting in more ‘‘smooth’’ decision boundaries. The amount of regularization is controlled by the α-parameter, which is set using a eight-fold cross-validation procedure on the training set. Generally all model selection is performed using K-fold cross-validation, regardless of the type of parameter that is to be determined (e.g. number of hidden nodes, size of the momentum term, etc.).

2.5.2. Neural network ensembles

A very common approach to increase the performance of an ANN classifier is to group many single ANNs into an ensemble of ANNs [5,9]. In the ensemble, the predictions of the individually trained networks are combined, usually with a simple averaging, to produce the ensemble prediction. This prediction is more accurate than the predictions of the individual ensemble members alone [5,9,14]. Combining the outputs of the ensemble members is only relevant when they disagree on some or several of the inputs. Techniques that have been used to create networks that disagree includes different initial weight setting, different network architectures or different subsets of the training set [14].

Currently, each ANN classifier in the WeAidU system consists of an ensemble of single multi-layer perceptrons. The necessary diversity of the individual members are created using different subsets of the training data. Typically the training set is randomly divided into K equally sized subsets that are used to train K different multi-layer perceptrons by excluding one subset each time. This procedure is repeated N times, which results in $K \times N$ ensemble members. The ensemble output is then simply the mean of the $K \times N$ ensemble members. For the experimental results presented in this paper $K = 3$ and $N = 6$ was used.

2.5.3. Diagnostic statements

One attractive quality of ANNs is that the predictions can be interpreted as Bayesian a posteriori probabilities [18]. This fact is used in the WeAidU system when post-processing the predictions of the classifiers to produce diagnostic statements that the physicians are familiar with. With four thresholds the output of the ANN classifier can be divided into five regions, corresponding to five different statements. For the diagnosis of infarction/ischemia, the following statements are used:

- absent,
- probably absent,
- equivocal,
- probably present,
- present.

Each threshold gives rise to a specific sensitivity and specificity (that can be estimated from the test set), which in turn corresponds to a transition between two statements. The thresholds should be selected such that clinically acceptable values for the corresponding sensitivity and specificity pairs are obtained.

2.6. Other statistical models

The performance of the ANN classifiers is compared with two other classification techniques, namely the k-nearest-neighbour (KNN) classification rule [2] and a method using logistic discrimination [1]. Both these methods provides Bayesian a posteriori probabilities. For the KNN method this is estimated using,

$$P(\mathscr{C}_{\mathrm{k}}|\boldsymbol{x}) = \frac{N_{\mathrm{k}}}{K} \tag{2}$$

where N_k is the number of data points that belongs to class $\mathscr{C}_k$ out of the possible k-nearest-neighbours of data point $\boldsymbol{x}$. The value of K was determined using cross-validation, similar to way parameters were set for the multi-layer perceptrons.

Logistic discrimination is implemented as simple perceptron with a logistic output activation function. The output $y(\boldsymbol{x})$, given the input $\boldsymbol{x}$, is computed as,

$$y(\boldsymbol{x}) = \frac{1}{1 + \exp(-\boldsymbol{w}^T\boldsymbol{x} - w_0)} \tag{3}$$

The weight vector $\boldsymbol{w}$ and the threshold w_0 are determined by minimizing the same cross-entropy error function as for the multi-layer perceptron.

2.7. Confidence levels and P-values

The area under the receiver operating characteristics (ROC) curve is used in this paper as a measure of performance. The 95% confidence levels can be estimated using a bootstrap method [21]. In short, re-sampling (with replacement) of the test set is used to construct B new ROC curves and their corresponding areas. The obtained distribution of ROC areas is used to calculate the 95% confidence levels. For more details see [21]. A value of $B = 50{,}000$ was used in this paper for the confidence levels presented in the result section. In principle the performance, and the accompanying confidence levels, are biased on our choice of training/test set split. However, the relatively large data set used will make this bias small.

The significance of the difference between two obtained ROC areas, for the same classification problem, was calculated using a permutation test (see e.g. [21]). The test is performed by repeatedly and randomly permuting the cases in the two lists. For each permutation the difference of the two resulting areas were calculated (test statistic). The evidence against the null hypothesis, of no difference between the two original ROC areas, was given by the fraction of area differences of the test statistic larger than the actual difference. The obtained P-values are used to find statistical significant differences between the performance of the three classification methods used in this paper.

3. Results

In this section, we show the performance of the ANN classifiers in the WeAidU system in terms of areas under the ROC curve. These results are also compared to the ones obtained by logistic discrimination and the k-nearest-neighbour method.

A visual presentation of some of the functionality of the WeAidU system is given at the end of this section.

3.1. Test set performance

The total study population of 1320 patients was randomly divided into one test set of 400 patients and one training set with the remaining 920 patients. For each classification task three models (ANN ensemble, logistic discrimination and k-nearest-neighbour) were

Table 1
The result of the ANN ensemble, logistic discrimination and the k-nearest-neighbour method, in terms of areas under the ROC curve

Territory	ANN ensemble	Logistic discrimination	k-nearest-neighbour
Detection of infarction			
Anterior	0.92 (0.88, 0.96)	0.89 (0.85, 0.94)	0.88 (0.83, 0.93)
Inferior	0.94 (0.92, 0.97)	0.94 (0.92, 0.97)	0.93 (0.90, 0.95)
Apical	0.93 (0.89, 0.97)	0.94 (0.91, 0.98)	0.93 (0.89, 0.97)
Septal	0.91 (0.87, 0.95)	0.88 (0.84, 0.93)	0.90 (0.87, 0.95)
Lateral	0.96 (0.94, 0.99)	0.96 (0.94, 0.99)	0.93 (0.89, 0.98)
Mean values	0.93	0.92	0.91
Detection of ischemia			
Anterior	0.87 (0.82, 0.92)	0.86 (0.81, 0.92)	0.77 (0.71, 0.83)
Inferior	0.83 (0.79, 0.88)	0.80 (0.75, 0.85)	0.76 (0.72, 0.81)
Apical	0.85 (0.79, 0.91)	0.83 (0.76, 0.90)	0.75 (0.67, 0.83)
Septal	0.85 (0.79, 0.90)	0.84 (0.78, 0.90)	0.78 (0.71, 0.84)
Lateral	0.87 (0.83, 0.93)	0.84 (0.78, 0.91)	0.86 (0.81, 0.92)
Mean values	0.86	0.83	0.78

The values in the brackets are 95% confidence levels estimated by a bootstrap method.

constructed using the training set. All model selection were performed using this training set. The performance presented below was measured on the test set. Table 1 lists the areas under the ROC curves for the 10 different classification tasks, namely detection of infarction and ischemia in five territories (anterior, inferior, apical, septal and lateral) of the heart. Generally the detection of infarction is easier than the ischemia detection as can be seen in Table 1.

Comparing the different classification methods we see a slight advantage using the ANN ensemble. The difference is larger for the detection of ischemia than it is for the infarction detection (comparing mean values). A deeper statistical comparison shows that there are statistical significant differences between the ANN ensemble and linear discrimination, for the infarction detection, in the following regions: anterior ($P = 0.014$) and septal ($P < 0.001$). The corresponding significant differences for the ischemia detection are: inferior ($P = 0.016$) and lateral ($P = 0.015$).

These results indicate, at least for the study population used in this paper, a small increase of the performance using ANN ensembles compared to linear discrimination (as implemented in this study).

3.2. Presentation of the WeAidU system

3.2.1. The client–server setup

The functionality of the WeAidU system is based on the client–server paradigm. The client consists of a program running on the physicians workstation, presenting the interface of the decision support. The server is a dedicated computer currently located at Lund University, Sweden.

The client and server programs are written in Java and the communication between them is handled by remote method invocation (RMI). RMI is developed by SUN and is the glue in a distributed object system such as a client–server application running on two different Virtual Machines. The choice of using Java as the implementation language enables multi-platform independence.

The user operating the WeAidU client program can send image data stored as files on the local hard disk or importing images directly into the client using the DICOM network protocol. Supported image file formates are: Interfile (version 3.3) and DICOM files (Spec 3.10.7), two of the most commonly used image formats for nuclear medicine images. The image data is sent to the server for analysis and other patient specific information is used for client display. The server stores information about the particular client request for feedback and usage statistics.

The response time depends to a great extent on the Internet-connection used by the client. Typical values lies within 1–20 s, after which an answer displaying the diagnostic advice will appear on the client screen (Fig. 2). The answer is combined with an appropriate feedback form. This user feedback is of vital importance for the future enhancement of the decision support system involved.

3.2.2. The WeAidU interface

The WeAidU system allows for the user to send image files stored on the local computer or to import images directly into the client using the DICOM network protocol. Each patient is identified by means of the ID, name and study data. A highlighted patient (study) can be send to the server for evaluation by pressing the *Evaluate* button (Fig. 2). It is important to note that only the image data is sent to the server for evaluation, since the WeAidU system (currently) only requires images data to make a diagnosis.

After having sent the image to the server the result will show up at the lower part of the interface (Fig. 2). There are two columns, one for the physician and one for the WeAidU interpretation. At present the system requires the physician to *actively* make his/her choice of interpretation before the WeAidU system shows its suggestion. This is an important detail since the system was developed as a decision *support* system where the physician still is the responsible person for the final diagnose.

There is also a possibility to look at the bulls-eye images that where analyzed. This may serve as a check for the physician that the images sent are the same as the ones displayed on the local system.

4. Discussion

The are four main reasons to believe that the use of computer-based decision support systems in the medical field have the potential for rapid expansion within the immediate future.

1. The development of interfaces can make computer-based decision support systems easily available to non-IT specialists such as the physicians, by employing state of the

art information transfer and assimilation techniques via WWW-browsers. These techniques make it possible to reach many physicians world-wide, as well as facilitating distribution of the latest version of the software.

2. Emerging techniques such as artificial neural networks makes it possible to develop more accurate decision support systems than would have been possible some years ago.
3. Vast medical databases in digital form can now be developed. These large information resources are important in order to create and control high quality decision support systems.
4. The infrastructure required for a widespread introduction of intelligent information systems, i.e. physicians workstations, networking, common standards for data-exchange, data storage, has been put in place in many hospitals, clinics, and surgeries, even in remote areas. Thereby, patient data, history, findings from physical examination as well as data from laboratory tests and diagnostic imaging can easily by used as input to an intelligent system.

The first item is very important, because the decision support system will not be accepted in the daily clinical practice if it requires many manual steps to operate or is too time-consuming. The accuracy is also an important ingredient for a working decision support system. Artificial neural networks is a very promising new technology for healthcare in general [13]. The use of this technique in the WeAidU system is motivated by several studies [3,4,10,16] showing the potential use of ANNs in connection with myocardial scintigraphy.

We believe the WeAidU system has potential for a successful decision support system that can operate in the daily clinical practice [15].

4.1. Clinical evaluation

How acceptable is the system to clinicians and how much do they trust the results? To answer these questions we invited 31 hospitals in Sweden to test the WeAidU system. We wanted to include different types of users, for example hospitals with >500 heart scintigrams per year, acquired by experienced personnel as well as small clinics performing heart scans only 1 day a week (<200 images per year). Seventeen of the 31 invited hospitals took part in the test of WeAidU. The remaining 14 hospitals did not participate for one of following three reasons: their gamma camera workstation was not connected to the Internet; the workstation software could not export images in standard formats; or due to the workload they did not have time to be involved in the test.

The participating hospitals used different gamma cameras, workstations and software. Some users installed the WeAidU system on a dedicated PC while others used their general gamma camera workstation. The fact that they used different hardware and software made it possible for us to test the generic nature of the system. The person technically responsible at each nuclear medicine department generally did the installation. Their impression was that it was easy to download the WeAidU client software and to install it on their PC or workstation. The physicians who used the WeAidU system, generally found it easy to use and of value in their clinical practice.

4.2. Future development

To fully assess the clinical importance of the WeAidU system proper clinical trials are needed that compare the clinical outcome for the patient with or without decision support.

Another issue is the safety and robustness of the ANNs involved in the interpretations of the myocardial perfusion images. Further development is needed to fully handle detection of outliers (e.g. bad quality images) and to assess measures of confidence for each classification performed. Similar methods have previously been developed by Holst et al. [7,8] for the classification of electrocardiograms.

Furthermore, the WeAidU system is constructed to handle different types of decision support, not only the interpretation of heart images. A similar system for the interpretation of lung images is currently being developed.

Acknowledgements

This study was supported by grants from the Swedish Foundation for Strategic Research, the Swedish Medical Research Council (K2001-27X-09893-10B) and the Medical faculty at Lund University.

References

[1] Anderson JA. Logistic discrimination. In: Krishnaiah PR, Kanal LN, editors. Classification, pattern recognition and reduction of dimensionality. Handbook of statistics, vol. 2. Amsterdam: North-Holland; 1982. p. 169–91.

[2] Bishop CM. Neural networks for pattern recognition. Oxford, UK: Oxford University Press; 1995. p. 187–9.

[3] Fujita H, Katafuchi T, Uehara T, Nishimura T. Application of artificial neural network to computer-aided diagnosis of coronary artery disease in myocardial SPECT bull's-eye images. J Nucl Med 1992; 33:272–6.

[4] Hamilton D, Riley PJ, Miola UJ, Amro AA. A feed forward neural network for classification of bull's-eye myocardial perfusion images. Eur J Nucl Med 1995;22:108–15.

[5] Hansen LK, Salamon P. Neural network ensembles. IEEE Trans Pattern Anal Machine Intell 1990; 12:993–1001.

[6] Hanson SJ, Pratt LY. Comparing biases for minimal network construction with back-propagation. In: Touretzky DS, editor. Advances in neural information processing systems. San Meteo, CA: Morgan Kaufmann; 1989. p. 177–85.

[7] Holst H, Ohlsson M, Peterson C, Edenbrandt L. A confident decision support system for interpreting electrocardiograms. Clin Physiol 1999;19:410–8.

[8] Holst H, Ohlsson M, Peterson C, Edenbrandt L. Intelligent computer reporting "lack of experience"—a confidence measure for decision support systems. Clin Physiol 1998;18:139–48.

[9] Krogh A, Vedelsby J. Neural network ensembles, cross validation, and active learning. In: Tesauro G, Touretzky D, Leen T, editors. Advances in neural information processing systems 2. San Mateo, CA: Morgan Kaufman; 1989. p. 650–9.

[10] Lindahl D, Palmer J, Ohlsson M, Peterson C, Lundin A, Edenbrandt L. Automated interpretation of myocardial SPECT perfusion images using artificial neural networks. J Nucl Med 1997;38:1870–5.

[11] Lindahl D, Lanke J, Lundin A, Palmer J, Edenbrandt L. Computer-based decision support system with improved classifications of myocardial bull's-eye scintigrams. J Nucl Med 1999;40:1–7.

[12] Lindahl D, Toft J, Hesse B, Palmer J, Lundin S, Ali A, et al. Scandinavian test of artificial neural network for classification of myocardial perfusion images. Clin Physiol 2000;20:253–61.

[13] Lisboa PJG, Ifeachor EC, Szczepaniak PS, editors. Artificial neural networks in biomedicine. London: Springer-Verlag; 2000.

[14] Opitz D, Maclin R. Popular ensemble methods: an empirical study. J Art Intell Res 1999;11:169–98.

[15] Ohren I. Evaluation of usability and usefulness of the WeAidU system as it is evaluated by the users, NST report 12-2002. Norway: Norwegian Centre for Telemedicine; 2002.

[16] Porenta G, Dorffner G, Kundrat S, Petta P, Duit-Schedlmayer J, Sochor H. Automated interpretation of planar thallium-201-dipyridamole stress-redistribution scintigrams using artificial neural networks. J Nucl Med 1994;35:2041–7.

[17] Press WH, Teukolsky SA, Vetterling WT, Flammery BP. Numerical recipies in C. 2nd ed. Cambridge: Cambridge University Press; 1995. p. 521–5.

[18] Richard M, Lippmann R. Neural networks estimate of Bayesian a posteriori probabilities. Neural Comput 1991;3:461–83.

[19] Rumelhart DE, McClelland JL, editors. Parallel distributed processing, vol. 1. Cambridge, MA: MIT Press; 1986. p. 318–62.

[20] Rvgnvaldsson T. On Langevin updating in multilayer perceptrons. Neural Comput 1994;6:916–26.

[21] Wehrens R, Putter H, Buydens LMC. The bootstrap: a tutorial. Chemom Intell Lab Syst 2000;54:35–52.