European Journal of
Nuclear
Medicine

Original article

A feed forward neural network for classification of bull's-eye myocardial perfusion images

David Hamilton[1], Peter J. Riley[1], Ueber J. Miola[1], Ahmed A. Amro[2]

[1] Department of Medical Physics, Clinical and Bio-Engineering, Riyadh Al Kharj Hospital Programme, Riyadh, Kingdom of Saudi Arabia
[2] Armed Forces Cardiac Centre, Riyadh, Kingdom of Saudi Arabia

Received 15 August and in revised form 3 October 1994

Abstract. Identification of hypoperfused areas in myocardial perfusion single-photon emission tomography studies can be aided by bull's-eye representation of raw counts, lesion extent and lesion severity, the latter two being produced by comparison of the raw bull's-eye data with a normal data base. An artificial intelligence technique which is presently becoming widely popular and which is particularly suitable for pattern recognition is that of artificial neural network. We have studied the ability of feed forward neural networks to extract patterns from bull's-eye data by assessing their capability to predict lesion presence without direct comparison with a normal data base. Studies were undertaken on both simulation data and on real stress-rest data obtained from 410 male patients undergoing routine thallium-201 myocardial perfusion scintigraphy. The ability of trained neural networks to predict lesion presence was quantified by calculating the areas under receiver operating characteristic curves. Figures as high as 0.96 for non-preclassified patient data were obtained, corresponding to an accuracy of 92%. The results demonstrate that neural networks can accurately classify patterns from bull's-eye myocardial perfusion images and detect the presence of hypoperfused areas without the need for comparison with a normal data base. Preliminary work suggests that this technique could be used to study perfusion patterns in the myocardium and their correlation with clinical parameters.

Key words: Myocardial perfusion – Single-photon emission tomography – Bull's-eye images – Neural network

Eur J Nucl Med (1995) 22:108–115

Introduction

Bull's-eye representation of single-photon emission tomographic (SPET) reconstructions and subsequent comparison with a normal data base is a commonly used method for aiding in the identification of hypoperfused areas in myocardial perfusion studies [1, 2] and has been shown to increase the diagnostic accuracy of the investigation significantly [3]. The technique involves the production of three parametric images comprising a map of the "raw" myocardial perfusion count distribution together with "extent" and "severity" maps of any detected lesions.

Another analysis technique which is presently becoming widely popular, and which is particularly suitable for the pattern classification necessary in work such as this, is that involving artificial neural networks (ANNs). These form a branch of artificial intelligence which dates back to 1949. Development of neural computing techniques since the middle 1980s has been extremely rapid [4, 5], with the current decade witnessing an explosive revival of research [6, 7] due in large part to the availability of low-cost, high-speed digital computers on which to simulate ANNs [6]. Most of today's research into ANN applications is undertaken on conventional computers using software simulations with hardware accelerator boards [8].

ANNs are systems which are trained to recognise similarities in patterns. They learn by example. Contributing to their popularity is the possibility that, because the pattern need not be defined, they may provide more information than existing techniques.

There are numerous types of ANN structures and learning strategies. One of the more straightforward designs is the feed forward neural network (FFNN). This is a network of neural nodes arranged in input, hidden and output layers which are fully connected, as illustrated in Fig. 1. Their operation is such that a data pattern, applied at the nodes of the input layer, will produce a pattern at the nodes of the output layer, determined by the weights of the connection functions. They can be trained by the application of a number of input-output pairs, constituting a training group, and the adjustment of the connection weights in accordance with learning rules until a required output pattern is achieved. The application of a complete training group constitutes an epoch.

Correspondence to: D. Hamilton, Department of Medical Physics, Clinical and Bio-Engineering, Military Hospital, PO Box 7897, Riyadh 11159, Kingdom of Saudi Arabia

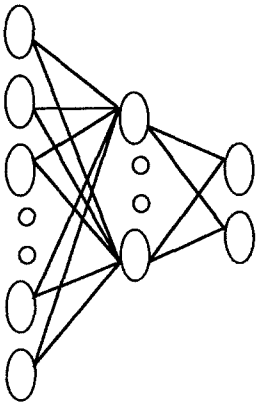

Fig. 1. A three-layer ANN showing the nodes as ellipses and the connecting weights as lines. The input, hidden and output layers are shown from *left to right*, the number of nodes in the hidden layer being defined as the geometric mean of the input and output nodes

During learning the training group is applied repetitively, the example inputs are presented to the network and the resultant and desired outputs are compared. An error term is then calculated which is used to alter network weights. Training continues until the network error is reduced below an output threshold, typically 0.1, termed convergence [8].

Four learning rules were available for this study, comprising: standard back projection, stochastic back projection, weigend weight eliminator and quick prop. The standard back projection assesses the gradient descent of an error surface. The errors are summed for every pattern in an epoch and then the weight changes are made based on this sum. Stochastic back propagation implements the weight changes after each pattern presentation. The weigend weight eliminator is similar to stochastic back propagation but also attempts to decrease the values of less important weights during training. Quick prop starts by using standard back propagation but then attempts to calculate the position of the error function minimum and jump straight to it.

The accuracy of the trained ANNs can be tested using new input patterns, constituting a testing group, and comparing the expected output pattern with that produced (guessed) by the ANN. One of the features of ANNs is their ability to generalise. This ability to learn and to generalise means that neural networks have the potential for solving image processing problems which are not readily tractable using rule-based conventional classifiers [7].

A number of articles have been written describing the application of ANNs in the medical field [7–14]. Specific examples in the analysis of non-imaging data have been discussed [15–21]. Those described in the imaging field sometimes rely on preclassification of the image pattern into numbers or symbols for input into the neural network, either by human interpretation [22–27] or by computer [28–30]. Thus neural network applications are starting to appear in the radiology literature and have even been used instead of conventional computer programs to reconstruct SPET images [31]. Few successful examples where the image is treated as an input to the network exist, however. One reason is the large volume of data in such images. Specific examples in imaging where preclassification is not undertaken tend to have restricted input data, either by using a very limited amount of information in a small coarse matrix [10, 32] or by human prelocalisation of a simulated lesion on SPET [33, 34]. Nuclear images, however (especially SPET), provide a good basis for testing the concept [35] and work of this nature is now starting to appear [36].

In terms of nuclear cardiology, a recent article reviewed ANN application to thallium-201 planar data. Best performance yielded an overall accuracy, compared to visual interpretation, of 88.7% on stress data. The best results were obtained for predicting the presence or absence of a defect, giving a sensitivity of 93% and a specificity of 83%. The same network trained by comparison with angiography gave only a 67% agreement, reflecting the fact that the input data are derived from a physiological variable and the output data are derived from an anatomical variable [37].

In terms of SPET bull's-eye myocardial perfusion images, of particular interest to the studies presented here, recent work has described image interpretation using lesion "extent" images as input to the ANN and comparison with angiographic lesion localisation to within the three main coronary arterial territories. This work gave an overall performance of 77% compared with human interpretation [35]. A sensitivity of 84% was achieved by comparing the ANN interpretation of clinical myocardial perfusion investigations with angiography [38]. Identification of abnormal vessels in single-vessel disease resulted in 100% accuracy, whereas in multiple-vessel disease the accuracy dropped to 80% [37].

We have previously presented "work in progress" [39, 40] which demonstrated the ability of FFNNs to discriminate ischaemia and infarction in the coronary vessel regions of the LAD, RCA and LCx with both a sensitivity and a specificity of 84% for mixed multivessel disease. When this "extents" FFNN was cascaded with our preliminary 93% accuracy detection FFNN, we obtained an overall "extents" accuracy of 91%.

Currently there are few theories governing the best way of constructing a network for a given problem and networks must be designed by trial and error. Careful pre-processing of the input data can dramatically reduce the number of input nodes, and so reduce the network size. Conversely, poor pre-processing can remove information required by the network to converge and so increase the training overhead. Increasing the number of input neurons may make the network learning task easier, but will also increase the number of network weights which must be set. An optimum training set is thought to be one containing many more examples than the number of network weights, and so coding strategies must not

increase the number of weights significantly if they are to be effective. Once the input and output layers have been defined, the number and size of the hidden layers must also be specified. Too few hidden nodes will inhibit learning, too many will allow the network to learn each training pattern individually and no generalisation will take place. Again, no real theories exist as to the number of nodes most suited to a problem of any given complexity [7]. Similarly, the total error allowed or tolerated in the training phase affects the performance of the network. The network can also be overtrained by the specification of a small total error in the training phase [36].

This report describes a series of studies undertaken to evaluate the feasibility of using FFNNs to aid in the interpretation of SPET myocardial perfusion images, using "raw" bull's-eye images without direct comparison to normal data. This has been done using raw count data as input and a determination of normality as output determined by comparison with the normal data base. The latter thus provides a consistent pattern with which to compare the behaviour of the neural network, which was studied by varying the learning rule, termination error, architecture, number of inputs and number of data sets in the training group.

Materials and methods

The ANNs were simulated on a digital computer using a commercial package (NeuralDesk, supplied by Neural Computer Sciences) comprising software running under Microsoft Windows on a 486DX platform and incorporating an accelerator card (NeuSprint).

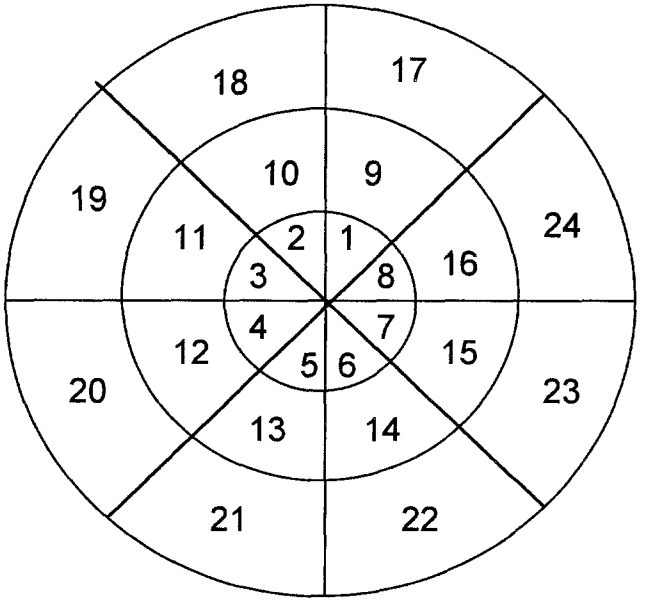

Fig. 2. The matrix imposed on the bull's-eye representation of myocardial perfusion scintigraphic data. Shown are the 24 input segments used in this study arranged anti-clockwise and concentrically from the apex, in the middle, to the base, at the periphery

The input data to the FFNN comprised raw counts from bull's-eye images produced by nuclear medicine acquisition and analysis software on a matrix of 15×40 pixels. This matrix was coarsened to 3×8 pixels as shown in Fig. 2 by summing counts over 5×5 pixel squares. It was felt that 24 values, rather than 600, would be the maximum that could practically be used to represent each bull's-eye image in this particular study. A subset of these 24 segments was also considered, comprising the eight segments of the middle circle. The input layer of the FFNN was designed so that each node would represent one pixel of the coarse matrix. The output layer consisted of a single node with a value coded as 0.99 for an abnormal state (lesion present anywhere in the myocardium) or coded as 0.01 for a normal state (no lesion present). Unless specifically determined otherwise, the architecture of the network was designed automatically by the software, which defined the number of nodes in the hidden layer as 5 when 24 input nodes were used and as 4 when eight input nodes were used.

The presence or absence of a lesion in a particular segment was defined using consistent criteria by coding as abnormal any segment that contained a raw count which was below 2.5 standard deviations (SD) of the expected mean count, obtained from a normal data base. This is the criterion used by Emory University [3] and was implemented with the normal data base supplied by IGE on their Star family of nuclear medicine computers. This data base is supplied on a matrix of 15×40 pixels and was coarsened for this study to a matrix of 3×8 to be consistent with the experimental data. As is recommended for ROC studies [41], equal numbers of abnormal and normal data sets were always used in each training group to avoid learning bias.

Both simulation and patient studies were undertaken to evaluate the ability of FFNNs to determine whether any segments in the myocardial perfusion image were below the threshold of normality.

Simulation studies. Simulation studies were undertaken in order to evaluate the behaviour of the FFNNs under different conditions using data that could be well controlled. The normal data base for stress studies was used as the basis for these simulation studies. These data are available as mean counts and standard deviation values and these were used to generate simulated patient data sets by combination with a set of normally distributed random numbers. Abnormally low counts in single segments were imposed on alternate data sets, thus producing alternate normal and abnormal data sets. The criterion of an abnormal segment count was one that fulfilled the Emory criteria of being more than 2.5 SD below the mean count value. These abnormally low segment counts were calculated by combining the mean count and standard deviation values from the normal data base with a set of random numbers uniformly distributed between 2.5 and 5.0. A population of data sets was constructed in this way, each set comprising 24 segments, with segment number 14 set low in both the training and testing groups. Segment 14 was chosen to be consistent with patient data presented later.

The following studies were undertaken using: the four learning rules of standard back projection, stochastic back projection, weigend weight eliminator and quick prop; terminating conditions of both a maximum error of 0.1 and an average error of 0.1; and 100 testing data sets. The FFNN parameters are summarised in Table 1. In *study 1* the variation in the number of training sets, from 60 down to 10, was considered, to evaluate the required number for optimal training. In *study 2* the value of the maximum error used as the terminating condition was varied from 0.05 to 0.4. In *study 3* the number of nodes in the hidden layer was varied from 3 to 12, including a consideration of two and three hidden layers of

Table 1. FFNN parameters used in the simulation study

Study	Training sets	Threshold	Hidden nodes	Bad segment training/testing
1	10–60	0.1	5	14/14
2	60	0.05–0.4	5	14/14
3	60	0.1	3–12, 5×5, 5×5×5	14/14
4	60	0.1	5	14/17
5	60	0.1	5	14, 17/17
6	60	0.1	5	14, 17/14, 17

five nodes each. In *study 4* the testing group was altered so that a different segment, number 17, was set low in each abnormal data set. In *study 5* one segment, number 17, was set low in each abnormal data set in the testing group. In *study 6* two segments, numbers 14 and 17, were set low in each abnormal data set in the testing group.

Patient studies. The data sets in this part of the study were constructed using images obtained from 410 male patients who underwent routine stress and rest ^{201}Tl SPET myocardial perfusion investigations. Male patients only were used because they formed the majority of referrals in the ratio of approximately 8:1. Male and female data were not mixed because the normal data bases for each are different and additional variability would have been introduced into the technique. Currently, we do not have enough data available from female patients to undertake a study for that group alone. The data were acquired using a step and shoot SPET acquisition protocol of 32 40-s frames on a 64×64 matrix, without zoom, over a circular orbit. The stress study was acquired in the morning and the rest study 4 h later. Reconstruction was undertaken using a Hanning prefilter with critical frequency of 0.82 cm^{-1} and a ramp back-projection filter. There was no correction for attenuation. Bull's-eye images were constructed according to the technique used by Emory University [3].

Each image segment was determined as being normal or abnormal by comparing the raw data counts with the counts calculated from the normal data base at the same matrix size. The normal data were normalised to the maximum count in each patient data set and abnormal segments were determined as those segments where the counts fell below 2.5 SD of the normal data mean count. As well as data groups comprising 24 input values, a subset was also considered comprising the eight segments from the middle circle shown in Fig. 2.

The first study involved data sets from both stress and rest data in which only one segment was low in each abnormal data set. The most numerous segment of this type was number 14 in the 24-segment group corresponding to number 6 in the 8-segment group. The networks were trained using all four learning rules and terminating conditions of both 0.1 maximum error and 0.1 average error. A jack-knife method was used to allocate the data sets between the training and testing groups [22]; that is, two-thirds of the data sets in each study population, selected randomly, were used for training, while the remaining one-third were used for testing. The numbers are shown in Table 2.

It was not possible to study the behaviour of specific multiple lesions in each abnormal data set, such as only two segments containing low counts, because sufficient relevant data were not available. Instead, for the second study, the non-preclassified raw data were considered. One hundred data sets were randomly allocated to the testing group. For the 24-input study, there was a maximum of 250 data sets in the training group for stress and 300 for delay. For the 8-input study, there was a maximum of 300 data sets in the training group for the stress study and 240 for the rest study. A data set was regarded as being abnormal if at least one segment was below 2.5 SD of the mean count.

Table 2. The number of data sets used in the training and testing sets in the following combinations. In each set, equal numbers of abnormal and normal data sets were incorporated

Data type	Number of input segments	Number of data sets	
		Training group	Testing group
Stress	24	28	12
	8	32	16
Rest	24	14	8
	8	30	14

Comparison of the above methods was made on the basis of receiver operating characteristics (ROC) analyses [42], in which the area under the ROC curve was used as the figure of merit [30, 43, 44]. The ROC curves were constructed by determining pairs of true-positive-ratio (TPR) and false-positive-ratio (FPR) values at different output thresholds on a scale of 0–1 [30]. Because many threshold values were available for the neural network performance, a non-parametric evaluation of the ROC area was used. Evaluating the threshold values of 0.1, 0.3, 0.5, 0.7, 0.9, 0.95 and 0.99 provided a good range of TPR and FPR values for plotting an ROC curve. The ROC curve for each trained network was constructed by calculating the sensitivity and specificity points, obtained from the output values of the tested network, for each value of the above thresholds in turn.

The area under this non-parametric ROC curve was estimated using the trapezoid rule [33]. The optimal parameters for the current neural network were determined by selecting those yielding the highest area under the ROC curves [22]. This varies from 0.5 when the ROC curve is based on guesswork to 1.0 when the decision process is totally accurate.

Results

Simulation studies

The first study involved data in which only one segment contained a low count and entailed varying the number of training sets for the learning rules of standard back projection, stochastic back projection, weigend weight eliminator and quick prop. All networks converged in less than 600 epochs and showed areas under the ROC

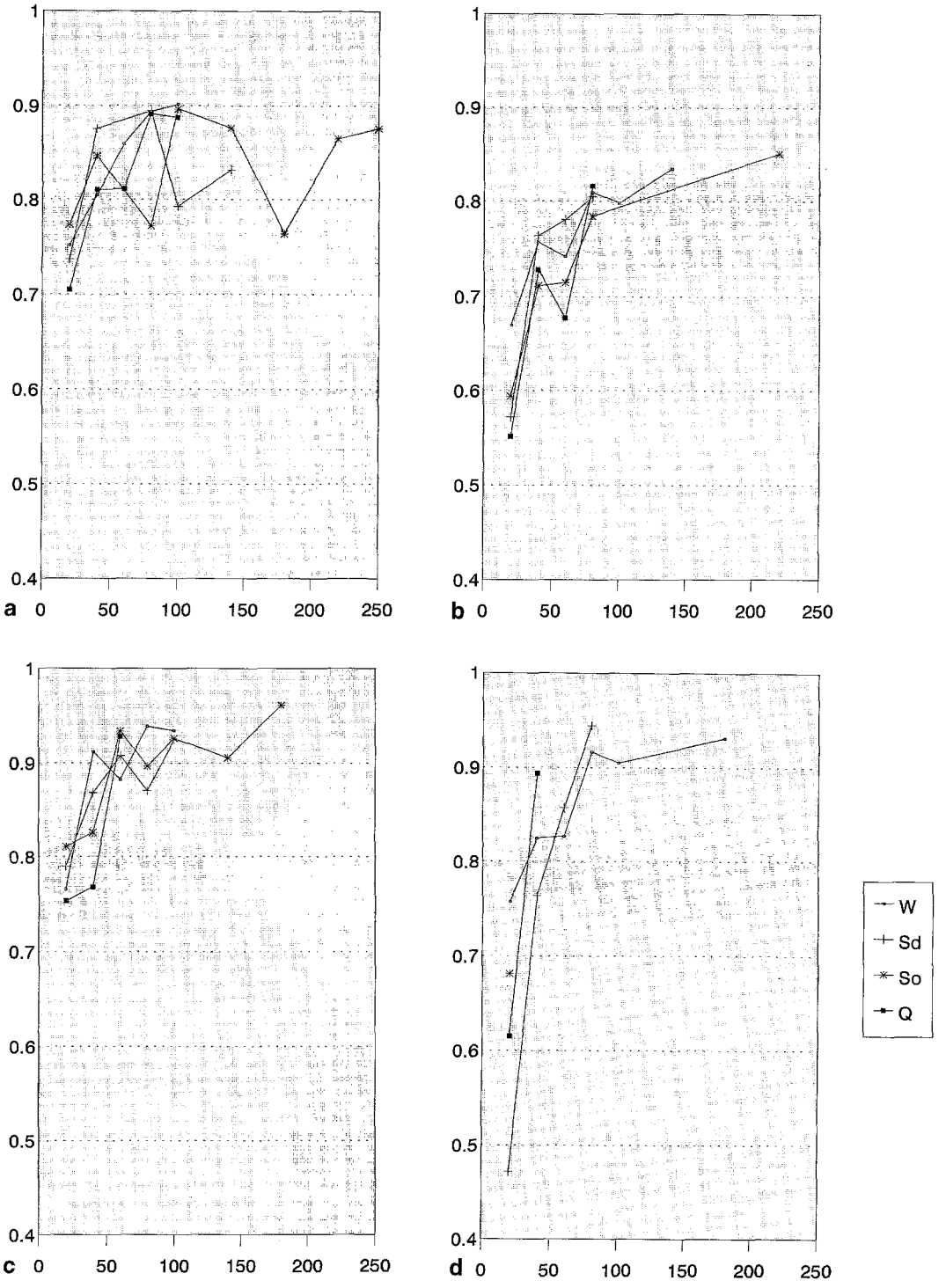

Fig. 3 a–d. The variation in area under the ROC curve (*ordinate*) for various numbers of data sets in the training group (*abscissa*). The figures show values for the four learning rules of: W = weigend weight eliminator, So = stochastic back projection, Sd = standard back projection and Q = quick prop. All used a terminating condition of a maximum error of 0.1. 24 input segments: **a** stress data, **b** rest data; 8 input segments: **c** stress data, **d** rest data

curves over 0.95. When more than 40 data sets were used in the training group, the areas were over 0.97. In the case of the three learning rules of standard back projection, stochastic back projection and weigend weight eliminator the areas started to decline when less than 40 data sets were used in the training group. The quick prop learning rule, however, maintained its performance down to ten data sets in the training group. Its performance was slightly better for the terminating condition of maximum error compared to the average error.

The results of the second study, which investigated the various terminating conditions showed areas under

the ROC curves of at least 0.99 in all cases. The number of training epochs required for convergence increased steadily from 120 to 2450, but there was very little change in areas under the ROC curves for the different terminating conditions. The results of the third study, investigating the effects of different network architectures defined by the number of hidden nodes, showed areas under the ROC curves of at least 0.99 in all cases. There was no significant change in the areas under the ROC curves, although the multiple layers did require fewer epochs to train, 160 for the 5×5 network compared to a minimum of 580 for the single hidden layer network.

In study 4, the network trained in 370 epochs using the standard back projection learning rule and when tested on data with segment 17 set low in each abnormal data set showed extremely poor results; the area under the ROC curve was 0.5. In study 5, the network trained in 1920 epochs using the stochastic back projection learning rule on data with both segments 14 and 17 set low in each abnormal data set. When tested on data with only segment 17 set low in each abnormal data set, it showed reasonably good results, with an area of 0.94. In study 6, the same network tested on data with both segments 14 and 17 set low in the same abnormal data set showed good results, with an area of 0.98.

Patient studies

The first patient study was undertaken on data in which only a single segment was low in each abnormal data set and was carried out for both stress and delay data and for 24 and 8 input segments. All of the four studies in this category trained well in between 140 and 7480 epochs and resulted in areas under the ROC curves of between 0.97 and 1.00 in the stress study and between 0.44 and 0.66 in the rest study for the 24 input segments, and of 0.98 in the stress study and 1.00 in the rest study for the 8 input segments.

The second patient study was undertaken on non-preclassified data for both 24 and 8 input segments. The training of these networks required various numbers of epochs, from less than 1000 to over 200 000, to converge. The results are shown in Fig. 3. These graphs show that, generally, accuracy improved in all studies as the number of data sets in the training group increased. Good accuracy was obtained for both stress and rest data with 24 input segments reaching a maximum of 0.88 for stress and 0.85 for delay. Improved accuracy was obtained, however, when 8 input segments were considered, reaching a maximum of 0.96 for stress data and 0.94 for rest data. In these cases, the figures translate to accuracies of 92% for stress and 93% for rest. It is evident that, with increasing numbers of data sets in the training groups, the FFNNs experienced more difficulty in converging and hence some data points in the graphs of Fig. 3 are absent, particularly at high numbers of data sets in the training groups.

Discussion

The simulation studies show that FFNNs can detect low count segments in well-controlled data very accurately. Networks trained on single lesions gave good results when over 40 data sets were included in the training group. One learning rule, quick prop, performed well down to ten data sets in the training group. A hidden layer comprising five nodes and a terminating condition of a maximum error of 0.1 gave the best results. Understandably, when the network was trained on data sets with a low count in one segment and tested on data sets with the low count in other segments the results were poor. However, accurate results were obtained when the number of low count segments in each abnormal data set was extended from one to two in each data set.

With patient data, except for the combination of rest data and 24 input segments, accurate results were again obtained when the training and testing groups comprised abnormal data sets containing a single lesion only. This was true for both 24 and 8 input segment groups. The poor accuracy of the 24 input segments rest data group was to be expected because the training group in this case comprised only 14 data sets whereas those of the other combinations contained between 28 and 32 data sets. For non-preclassified patient data which contained numerous different patterns of multiple low count segments, good accuracy was obtained for both stress and rest in the 24 segments but better accuracy was obtained when only 8 segments from the middle circle were considered, giving a maximum area under the ROC curve of 0.96. This accuracy was obtained without any processing of the data to exclude difficult cases or to make a more homogeneous group. This finding indicates that the FFNNs can consistently detect low count segments without the aid of a normal data base and suggests that an algorithm including more than one network working in concert might be useful [39, 40, 45].

The output used in these studies was based, effectively, on the "extent" images generated for each study. It was used to demonstrate the feasibility of the technique and cannot pretend to be using the method to its full potential. Out of necessity, a very coarse matrix was used. This has the effect of smoothing the observed perfusion pattern over the myocardium and subtle variations which might be represented on the fine matrix of 600 pixels are likely to be lost on this coarse matrix of 24 pixels. Results, although good, might well improve when facilities for handling larger amounts of data become more widely available. Our results confirm that the performance of an ANN is highly dependent on the composition of the data on which it is trained since the ability of a network to identify a pattern is directly related to the representation of that pattern in the data used to train the network [37, 45] and that a major problem is limitations in defect sampling. For the FFNN to generalise well to new cases, the data used in training must be representative of the

full population of data likely to be sampled, because each possible defect location, plus all the possible combinations of locations, must be included [37]. This was shown to be unlikely to happen when the number of positive cases of each type is less than 50. This is the main disadvantage of FFNNs, and implies that a very large patient population is required to provide a full set of lesions for training. Also, a network which is accurate in some cases could fail on data which are not represented by the training data.

The results demonstrate clearly that FFNNs can be trained to indicate, consistently, the clinical state of the myocardial perfusion using only "raw" bull's-eye images and without requiring individual direct comparison to normal data. The method, however, requires the use of a normal data base for both training and testing and is therefore not independent of the data. It has been used only to provide a comparison with conventional data base techniques to determine whether the FFNN performs equally well. It would be preferable to assess the performance of the FFNN by comparison with independent criteria. However, no such criteria are currently easily available.

Though neural network training is strictly a "supervised learning" process, the learning is essentially by example, with no guidance from the user as to the criteria to employ. The network is allowed to learn what it "believes" to be the most important discriminating features, and to weigh those features appropriately for best classification performance [30]. Thus, one of the features of FFNNs is that they extract any linking pattern between input and output codes themselves without requiring a separate pattern definition; the pattern need not be known other than by the FFNN.

A development with possibly great potential would be the extraction of patterns from the bull's-eye images without the constraint imposed by prior comparison with normal data. This could have important implications in that normal comparison has a constraining effect on the pattern extraction process and the FFNN technique allows this limiting factor to be removed. Since it is becoming apparent that a positive ^{201}Tl and a negative angiogram put patients into an "at risk" group [46–48], such correlations should be instructive in evaluating myocardial perfusion abnormalities at an early stage in their development. They could also include the investigation of links between the global myocardial perfusion pattern and a variety of distant clinical outcomes for which linking pattern definition would be difficult, e.g. the suitability of the heart for percutaneous transluminal coronary angioplasty/coronary artery bypass graft or the prediction of future coronary artery stenosis. This would be instead of restricting the analysis to the identification of hypoperfusion and comparison with angiography. This technique is especially powerful since, unlike the human brain, neurals can be cleared (randomised) of previous experiences and trained from scratch [37] to find new linking patterns.

Conclusion

The ability of FFNNs to extract patterns from bull's-eye data by assessing their capability to predict lesion presence without direct comparison with a normal data base were investigated. Studies were undertaken on both simulation data and on real stress-rest data obtained from 410 male patients undergoing routine ^{201}Tl myocardial perfusion scintigraphy. The ability of trained neural networks to predict lesion presence was quantified by calculating the areas under ROC curves. Figures as high as 0.96 for non-preclassified patient data were obtained. This translates, in this case, to an accuracy of 92%. The results demonstrate that neural networks can accurately extract patterns from bull's-eye myocardial perfusion images and detect the presence of hypoperfused areas without the need for comparison with a normal data base. It is postulated that this technique could be used to study perfusion patterns in the myocardium and their correlation with clinical parameters.

Acknowledgements. The authors would like to acknowledge the help and support of the staff of the Nuclear Medicine Section of the Department of Medical Physics, Clinical and Bio-Engineering in this work.

References

1. Caldwell JH, Williams DL, Harp GD, Stratton JR, Richie JL. Quantitation of size of relative myocardial perfusion defect by single-photon emission computed tomography. *Circulation* 1984; 70: 1048–1056.
2. Garcia EV, van Train K, Maddahi J, Prigent F, Friedman J, Areeda J, Waxman A, Berman DS. Quantification of rotational thallium-201 myocardial tomography. *J Nucl Med* 1985; 26: 17–26.
3. DePasquale EE, Nody AC, DePuey EG, Garcia EV, Pilcher G, Bredlau C, Roubin G, Gober A, Gruentzig A, D'Amato P, Berger H. Quantitative rotational thallium-201 tomography for identifying and localizing coronary artery disease. *Circulation* 1988; 77: 316–327.
4. Wasserman PD. *Neural computing. Theory and practice.* New York: Van Nostrand Reinhold, 1989.
5. Eberhart RC, Dobbins RW. *Neural network PC tools: a practical guide.* San Diego: Academic Press, 1990.
6. Clark JW. Neural network modelling. *Phys Med Biol* 1991; 36: 1259–1317.
7. Miller AS, Blott BH, Hames TK. Review of neural network applications in medical imaging and signal processing. *Med Biol Eng Comput* 1992; 30: 449–464.
8. Scott R. Artificial intelligence: its use in medical diagnosis. *J Nucl Med* 1993; 34: 510–514.
9. Boone JM, Gross GW, Greco-Hunt V. Neural networks in radiologic diagnosis. I. Introduction and illustration. *Invest Radiol* 1990; 25: 1012–1016.
10. Boone JM, Sigillito VG, Shaber GS. Neural networks in radiology: an introduction and evaluation in a signal detection task. *Med Phys* 1990; 17: 234–241.
11. Guerriere MRJ, Detsky AS. Neural networks: what are they? [editorial]. *Ann Intern Med* 1991; 115: 906–907.

12. Ezquerra N, Garcia E. Artificial intelligence gives computer new role as imaging problem-solver. *Diagn Imag* 1985; 7: 195–200.
13. Makhoul J. Artificial neural networks. *Invest Radiol* 1990; 25: 748–750.
14. Shufflebarger CM, Young W. What is a neural network? [editorial]. *Ann Emerg Med* 1992; 21: 1461–1462.
15. Allen J, Murray A. Development of a neural network screening aid for diagnosing lower limb peripheral vascular disease from photoelectric plethysmography pulse waveforms. *Physiol Meas* 1993; 14: 13–22.
16. Akay M. Noninvasive diagnosis of coronary artery disease using a neural network algorithm. *Biol Cybern* 1992; 67: 361–367.
17. Baxt WG. Use of an artificial neural network for data analysis in clinical decision making: the diagnosis of acute coronary occlusion. *Neural Computation* 1990; 2: 480–489.
18. Baxt WG. Use of an artificial neural network for the diagnosis of myocardial infarction. *Ann Intern Med* 1991; 115: 843–848.
19. Baxt WG. Analysis of the clinical variables driving decision in an artificial neural network trained to identify the presence of myocardial infarction. *Ann Emerg Med* 1992; 21: 1439–1444.
20. Bounds DG, Lloyd PJ, Mathew BG. A comparison of neural network and other pattern recognition approaches to the diagnosis of low back disorders. *Neural Networks* 1990; 3: 583–591.
21. Nelson TR, Boone JM. Visualization of myocardial activation sequences with neural network-based electrocardiographic localization: simulation results [abstract]. *RSNA* 1992; 56.
22. Asada N, Doi K, MacMahon H, Montner SM, Giger ML, Abe C, Wu Y. Potential usefulness of an artificial neural network for differential diagnosis of interstitial lung diseases: pilot study. *Radiology* 1990; 177: 857–860.
23. Fukuda H, Usuki N, Salwai S, Nakajima H, Miyamoto T, Inoue Y. Usefulness of an artificial neural network for assessing the ventricular size [abstract]. *RSNA* 1992; 56.
24. Gross GW, Boone JM, Greco-Hunt V, Greenberg B. Neural networks in radiologic diagnosis. II. Interpretation of neonatal chest radiographs. *Invest Radiol* 1990; 25: 1017–1023.
25. Piraino DW, Amartur SC, Richmond BJ, Schils JP, Thome JM, Belhobek GH, Schlucter MD. Application of an artificial neural network in radiographic diagnosis. *J Dig Imag* 1991; 4: 226–232.
26. Scott JA, Palmer EL. Neural network analysis of ventilation-perfusion lung scans. *Radiology* 1993; 186: 661–664.
27. Wu Y, Doi K, Giger ML, Metz CE, Zhang W. Detection of lung nodules on digital radiographs: comparison of artificial neural networks and discriminant analysis [abstract]. *RSNA* 1992; 56.
28. Chan KK, Hayrapetian AS, Lau CC, Lufkin R. Neural network segmentation of double-echo MR images [abstract]. *RSNA* 1992; 157.
29. Kim JH, Min BG, Han MC, Lee CW. Computer-assisted detection of lung nodules by using artificial neural net [abstract]. *RSNA* 1992; 56.
30. Kippenham JS, Barker WW, Pascal S, Nagel J, Duara R. Evaluation of a neural-network classifier for PET scans of normal and Alzheimer's disease subjects. *J Nucl Med* 1992; 33: 1459–1467.
31. Floyd CE, Bowsher JE, Munley MT, Tourassi GD, Baydush AH, Coleman RE. Neural network for quantitative reconstruction of SPECT images [abstract]. *J Nucl Med* 1991; 32: 936.
32. Garg S, Floyd CE. Neural network localization of pulmonary nodules on digital chest radiographs [abstract]. *RSNA* 1992; 157.
33. Floyd CE, Tourassi GD. An artificial neural network for lesion detection on single-photon emission computed tomographic images. *Invest Radiol* 1992; 27: 667–672.
34. Tourassi GD, Floyd CE, Coleman RE. Detection and localization of cold lesions on SPECT images with artificial neural networks [abstract]. *RSNA* 1992; 157.
35. 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: 272–276.
36. Chan KH, Johnson KA, Becker JA, Satlin A, Mendelson J, Garada B, Holman BL. A neural network classifier for cerebral perfusion imaging. *J Nucl Med* 1994; 35: 771–774.
37. Datz FL, Rosenberg C, Gabor FV, Christian PE, Gullberg GT, Ahluwalia R, Morton KA. The use of computer-assisted diagnosis in cardiac perfusion nuclear medicine studies: a review (part 3). *J Dig Imag* 1993; 6: 67–80.
38. Wang DC, Juni JE. Artificial neural network interpretation of cardiac stress thallium studies [abstract]. *RSNA* 1992; 283.
39. Hamilton D, Riley PJ, Miola UJ, Amro A. Neural network analysis of Tl-201 SPECT bullseye images [abstract]. Proc: 1st International Congress of Nuclear Cardiology, Cannes, France, April 1993.
40. Riley P, Hamilton D, Miola UJ, Amro A. Neural network discrimination of myocardial ischaemia and infarction in ^{201}Tl SPECT imaging [abstract]. *Br J Radiol* 1993; C75.
41. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978; 8: 283–298.
42. Metz CE. ROC methodology in radiologic imaging. *Invest Radiol* 1986; 21: 720–733.
43. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristics (ROC) curve. *Radiology* 1982; 143: 29–36.
44. Meistrell ML. Evaluation of neural network performance by receiver operating characteristic (ROC) analysis: examples from the biotechnology domain. *Comput Methods Programs Biomed* 1990; 32: 73–80.
45. Baxt WG. Improving the accuracy of an artificial neural network using multiple differently trained networks: *Neural Computation* 1992; 4: 772–780.
46. Mundler O, Grousset C, Birkui P, Pauchet M. Two years follow up of patients with "false positive" exercise or dipyridamole thallium 201 myocardial SPECT [abstract]. *Eur J Nucl Med* 1989; 15: 416.
47. Rosen SD, Camici PG. Syndrome X: radionuclide studies of myocardial perfusion in patients with chest pain and normal coronary arteriograms. *Eur J Nucl Med* 1992; 19: 311–314.
48. Solot G, Hermans J, Merlo P, Chaudron J-M, Luwaert R, Cheron P, Bodart F, Beauduin M. Correlation of ^{99}Tc^{m}-sestamibi SPECT with coronary angiography in general hospital practice. *Nucl Med Commun* 1993; 14: 23–29.