Artificial Intelligence in Medicine 16 (1999) 205–222

Artificial Intelligence in Medicine

An intelligent framework for the classification of the 12-lead ECG

C.D. Nugent [a,*], J.A.C. Webb [a], N.D. Black [a], G.T.H. Wright [a], M. McIntyre [b]

[a] *The Northern Ireland Bio-engineering Centre, School of Electrical and Mechanical Engineering, University of Ulster at Jordanstown, Shore Road, Newtownabbey, BT37 0QB, UK*

[b] *Regional Medical Cardiology Centre, Royal Victoria Hospital, Grosvenor Road, Belfast, UK*

Received 29 June 1998; received in revised form 12 October 1998; accepted 2 November 1998

Abstract

An intelligent framework has been proposed to classify an unknown 12-Lead electrocardiogram into one of a possible number of mutually exclusive and combined diagnostic classes. The framework segregates the classification problem into a number of bi-dimensional classification problems, requiring individual bi-group classifiers for each individual diagnostic class. The bi-group classifiers were generated employing Neural Networks (NN), combined with a combination framework containing an Evidential Reasoning framework to accommodate for any conflicting situations between the bi-group classifiers. A number of different feature selection techniques were investigated with the aim of generating the most appropriate input vector for the bi-group classifiers. It was found that by reducing the original input feature vector, the generalisation ability of the classifiers, when exposed to unseen data, was enhanced and subsequently this reduced the computational requirements of the network itself. The entire framework was compared with a conventional approach to NN classification and a rule based classification approach. The framework attained a significantly higher level of classification in comparison with the other methods; 80.0% compared with 66.7% for the rule based technique and 68.00% for the conventional neural approach.

Keywords: Computerised electrocardiography; Neural networks; Feature selection; Evidential reasoning

* Corresponding author. Tel.: +44-1232-368013; fax: +44-1232-366863.
E-mail address: cd.nugent@ulst.ac.uk (C.D. Nugent)

PII: S0933-3657(99)00006-8

1. Fundamentals of cardiac electrophysiology

Cardiac electrophysiology refers to the electrical phenomena associated with the physiological processes within the heart. The rhythmic contraction of the heart pumps blood around two parallel loops: the systemic circulation and the pulmonary circulation. The systemic circulation carries the freshly oxygenated blood throughout the body to vital organs, whilst the pulmonary circulation caries the deoxygenated blood from the systemic circulation back to the lungs for oxygenation. The heart is forced to pump blood (contract) by the spread of electrical currents throughout its muscle. Complex changes in ionic concentration across the cell membrane establish an extracellular potential field, exciting neighbouring cells and initiating propagation of electrical events from cell to cell. The body may be considered to be a thoracic volume conductor, a purely passive medium containing no electrical sources or sinks [3], therefore the potential fields established are carried to the body surface. At any instant of time it is possible to represent the electrical activity of the heart by a net equivalent current dipole, located at the electrical centre of the heart. The resistive load of the heart, or equivalent cardiac generator may be considered to be the thoracic medium. (The dipole concept is only an approximation to give a complete description of body surface potentials. Some of the potentials recorded may be attributed to non dipolar sources [16,18]). The amount of tissue activated, together with the relative speed and direction of the activation wavefront, determines the characteristic shape of the body surface waves. By placing electrodes on the body surface it is possible to detect the voltages of the cardiac currents. The most commonly employed method to record the electrical activity of the heart is the 12-lead electrocardiograph [4,8]. This technique records the sequence of electrical signals of the heart, referred to as the electrocardiogram (ECG) from various positions on the surface of the patient's body, in essence allowing the heart to be viewed from a number of different angles. Each component of the ECG is directly related to the spread of electrical currents through specific regions of the heart (Fig. 1). Thus sufficient information is available in these signals to enable diagnosis of a number of cardiac abnormalities. The P wave is representative of atrial depolarisation (cardiac stimulation), the QRS complex represents

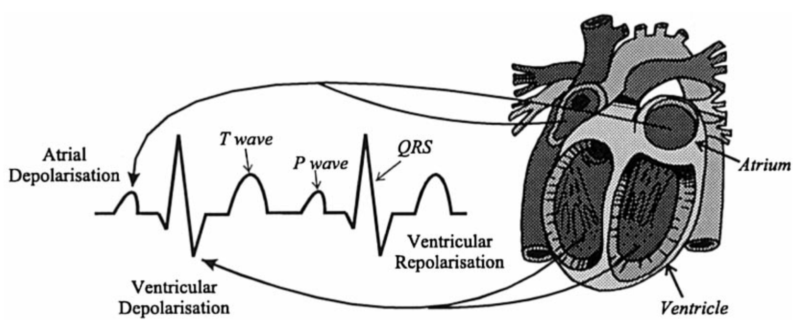

Fig. 1. Spread of Electrical Stimulus Throughout the Heart.

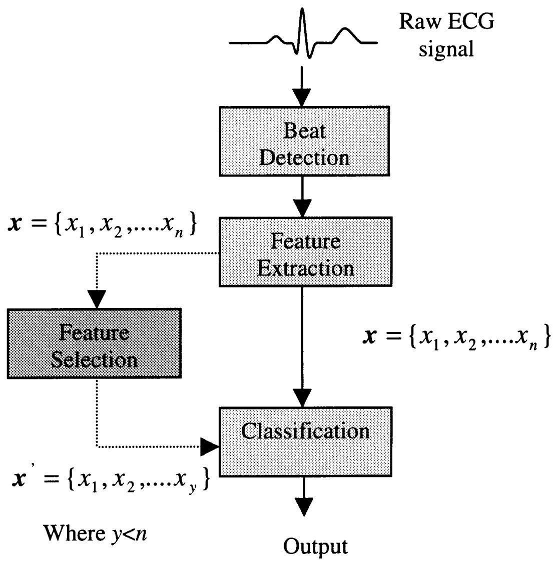

Fig. 2. Overview of functional modules in a typical Computerised Electrocardiographic system.

ventricular depolarisation and the T wave represents the return of the ventricles to their resting state (repolarisation). There is no visible waveform for atrial repolarisation as it is engulfed by ventricular depolarisation.

2. Computerised ECG classification

Since the advent of computerised electrocardiography in 1957 by Pipberger [20] research techniques in the field have proliferated. This was one of the earliest forms of the introduction of computers to medicine, simulating a similar medical decision making process to that adopted by clinicians. The goal of ECG classification is to classify the unknown recorded signal into one of a possible number of diagnostic classes, determining if the patient's cardiac condition is 'normal' and may remain untreated, or whether the patient exhibits any cardiac abnormalities and requires a form of treatment. The same electrical information, produced by the heart during each cardiac cycle, is examined by both human and computerised techniques. However, to some extent computerised approaches have the ability to produce more consistent results [23] and are able to consider more information from the ECG signal at higher speeds and in more complex permutations than the human interpreter. Various methodologies of automated diagnosis have been adopted, however the entire process can generally be subdivided into a number of disjoint processing modules: Beat Detection; Feature Extraction/Selection; and Classification (Fig. 2). The initial pre-processing module of Beat Detection aims to locate each cardiac cycle in each of the recording leads and insert reference markers indicating the beginning and end of each interwave component [13]. The algorithm is designed with two main objectives: firstly, the detector should provide reliable detection of each cardiac cycle in all recording leads and secondly, the temporal location of the reference points should be described accurately. The accuracy of detection of each cardiac cycle is of great importance since it contributes significantly to the overall classification result. The markers are subsequently processed by the Feature Extraction Module, where measurements are produced for wave amplitudes and durations. The collective term for the measurements produced is commonly referred to as the input feature vector, which is considered to describe the morphology of the current recorded signal. The module of feature selection is an optional stage, whereby the feature vector is reduced in size including only, from the classification viewpoint, what may be considered as the most relevant features required for discrimination. The classification module is the final stage in automated diagnosis. It examines the input feature vector and based on its algorithmic nature, produces a suggestive hypothesis.

2.1. Classification techniques

A multitude of computerised algorithms have been employed to accommodate the final process of classification of the ECG, based on the acquired digital information. Generally the techniques employed may be segregated into two categories. Firstly those based on a multivariate model, generally in the form of multivariate statistical techniques and Neural Networks (NN). Secondly, heuristic techniques employing rule based models. Heuristic classifiers include Decision Trees, Fuzzy Logic and Expert Systems, all simulating the reasoning of the human expert. Pipberger [17] referred to heuristic classifiers as First Generation programs and multivariate techniques as Second Generation programs. The labels assigned were not necessarily indicative of their time of occurrence but with respect to their complexity level. More recently, enhanced computing power, programming techniques and research experience perhaps herald the advent of 'Potential Third Generation Programs' [14].

The most pertinent problem with computerised classification, which in some respects may be attributed to its popularity of research, is the lack of standardised classification rules and clear definitions of what information should be considered during diagnosis [9,25]. This problem mainly effects rule based classifiers as they are dependent on the knowledge provided by the human expert, which in some cases may be based on trial and error. This problem of lack of standardisation may explain the recent popularity of applications of Artificial Intelligence where no rules are initially required and the human input is used to annotate a train and test database, around which the classifier adapts itself.

A second generation classifier is proposed to accommodate diagnosis of a subset of Category A statements, as defined by the Common Standards for Quantitative Electrocardiography Committee [24]. The diagnostic classes include Inferior Myocardial Infarction (IMI), Anterior Myocardial Infarction (AMI), a combination of Inferior and Anterior Myocardial Infarction (CMI), Left Ventricular Hypertrophy

(LVH), a combination of Hypertrophy and Infarctions (CMILVH) and finally Normal. The classification framework employs a family of bi-group NN, combined with a post-processing module containing an Artificial Intelligence framework based on the Demspter–Shafer Theory of Evidence, employed for error control. Various feature selection techniques have been investigated, aiming to obtain the optimal input feature vector for each individual bi-group classifier.

3. Bi-group classification proposal

To formally define a classification procedure, we begin by considering a set of N labelled samples, where x is an N-dimensional feature vector, $\boldsymbol{x} = [x_1,\ldots,x_N] \in R^N$, drawn from some underlying distribution $s(x)$, with class label l. For M class classification problems, the associated class for x may be expressed as $j \in \{1,\ldots,M\}$, where M is the total number of possible classes and the class label as $l \in \{0, 1\}^M$. The j^{th} element of l is one if x belongs to class j and the rest of the elements of l are zero.

Case 1: Conventional classifier for M class problem. Considering an unknown pattern $\boldsymbol{x}$, drawn from $s(x)$, the classifier itself may be considered as a function $u: R^N \rightarrow \{0, 1\}^M$, producing the required classification result.

Case 2: Classification employing bi-group classifiers. Alternatively the above M class classification problem can be represented by a family of bi-group classifiers with a range space $A = \bigcup_{i=1}^{M} A_i$ where $A_i \cap A_j = 0$, $i \neq j$. Each individual classifier may be considered to perform the function $BG_i: R^N \rightarrow A_i, 1 \leq i \leq M-1$ where $BG_i(\boldsymbol{x})$ equals 1 if x belongs to class i and zero otherwise. For $i = M$ and $\alpha \in A_M$, $\alpha \neq 0$ provided $BG_i(\boldsymbol{x}) = 0$, $\forall 1 \leq i \leq M-1$, $\boldsymbol{x} \in R^N$.

Example: Consider the pattern $\boldsymbol{x}$ with associated Class 2, drawn from $s(x)$ with $M = 3$.

Case 1: The classification function may be expressed as,

$$u(\boldsymbol{x}) \in \{0, 1\}^3$$

giving a final result of,

$$u(\boldsymbol{x}) = 010$$

Case 2: Each of the individual bi-group classifiers must be considered producing,

$$BG_i(\boldsymbol{x}) = 0,\ BG_2(\boldsymbol{x}) = 1,\ \therefore \alpha = 0$$

therefore presenting the final result of,

$$BG_i(\boldsymbol{x}) = 010$$

3.1. Bi-group NN

Each of the individual bi-group classifiers have been generated employing NN. Over the past decade, applications of NN have proliferated in many areas of

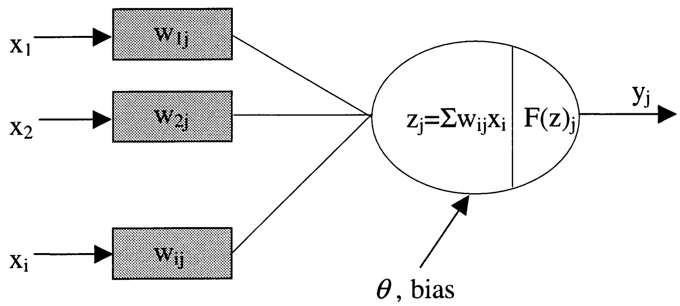

Fig. 3. Artificial neuron applying transformation of summed weighted inputs.

decision support. They have been applied to problems in areas of pattern recognition/classification, control and optimisation and have been particularly useful in cases where the problem to be solved is ill defined and development of an algorithmic solution is difficult. They have been applied in many fields of medicine [10] such as cancer prognosis, magnetic resonance imaging [1], diagnosis of low back disorders [2], detection of coronary artery disease [5] plus many more. In the past decade, the research in the field of ECG classification employing NN has been widespread and ever increasing. Their popularity to some extent may be attributed to their powerful parallel processing powers and adaptive nature, where no set of classification rules must be specified. A number of different architectures, along with training algorithms exist (see [7,11,12,22] for comprehensive overview), however in the field of ECG classification the most commonly employed NN has been seen to be the multi-layered perceptron, trained with back propagation.

In general terms a NN distinguishes itself from a conventional computer in that it is a parallel distributed non-linear computing network, mimicking the information processing structure of a biological neural system. Fig. 3 is representative of an 'artificial neuron' which aims to emulate the behaviour of the biological neuron.

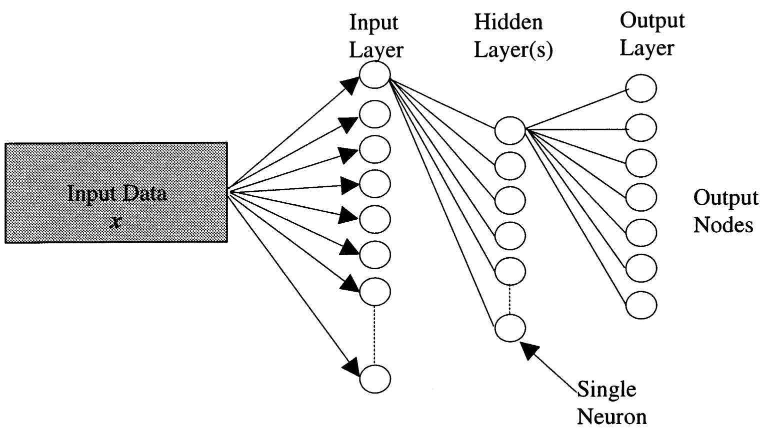

Fig. 4. Multi-layer perceptron consisting of an input layer, an output layer and one hidden layer.

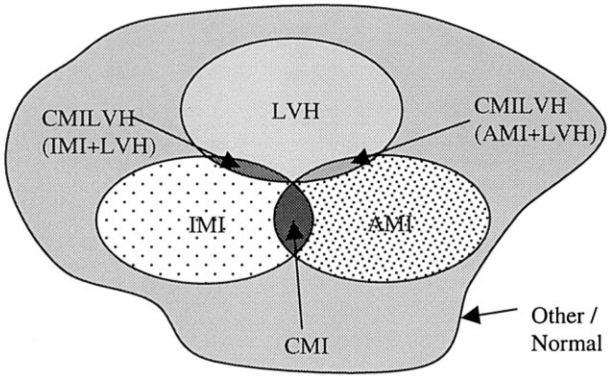

Fig. 5. Classification space to be accommodated by proposed framework.

The body of the neuron is represented by a weighted sum of the input signals $(z)j$ and the bias, followed by a linear or non-linear function $F(z)j$, referred to as the activation function. If neurons are grouped together in layers, with weighted synapses interconnecting only neurons in successive layers, the NN structure is generally referred to as a mulit-layer perceptron. A multi-layer perceptron consists of an input layer, an output layer, with one or more hidden layers [6], see Fig. 4. With sufficient numbers of synapse weights and number of internal layers, such a structure is capable of approximating any non-linear functional mapping with an arbitrary degree of accuracy. The application of NN involves a training and testing phase. During supervised training, patterns, along with the desired outputs, are presented to the network. A training algorithm is employed which adjusts the values of the internal weighting factors of the network by an amount proportional to the difference between the desired and actual output. The network when fully trained, as determined by the training algorithm is then exposed to a test set of data in order to evaluate the performance of the network on unseen data.

A bi-group NN is simply a multi-layer perceptron with a single node in the output layer. The single output reduces the classification problem to one in bi-dimensions, where either the presence or absence of a specific class is indicated. Reduction of the classification problem to bi-dimensions, subsequently reduces the feature space to one of also bi-dimensions, hence simplifying the feature selection process. With regard to training, the subsets selected from the initial training population for the individual classifiers allow more attention to be given to specific classes and hence increase the generalisation ability of the networks in comparison to a multi-output classifier.

3.2. Bi-group NN framework

The classification ability of the framework, as already mentioned, is based on a subset of Category A diagnostic statements. Fig. 5 is representative of the classification space. It is assumed at this point that two possible situations may arise. Firstly, the patient may belong to one of a possible three mutually exclusive regions i.e. IMI, AMI or LVH only. Secondly, due to the overlapping regions, a patient may exhibit a combination of different types of infarctions and hypertrophy or a combination of different types of infarctions. In the first instance it is possible to generate bi-group classifiers to identify the presence or absence of the classes under mutually exclusive like conditions. In the second instance, combinatorial like classifiers are required to identify the specialised regions of the overlapping classes. It would be possible to simply employ a family of three bi-group classifiers to segregate the space, assuming that under all possible scenarios, they are of a non-mutually exclusive nature. However by including the additional two bi-group combinatorial classifiers, an additional security measure is incorporated into the framework, avoiding misclassification and identification of possible conflicting situations between two mutually exclusive diagnostic classes. When the situation arises that the pattern $\boldsymbol{x}$ belongs to an overlapping region, two of the mutually exclusive like bi-group classifiers will be asserted, in conjunction with the appropriate combinatorial network. The inclusion of both methodologies provides sufficient information to indicate the presence of an overlapping classification region and helps prevent misclassification of non-overlapping regions caused by conflicting outputs of the mutually exclusive like networks. It is also assumed that if the unknown pattern $\boldsymbol{x}$ does not belong to any of the diagnostic classes specified, it is assigned to the Normal class. Fig. 6 is indicative of the internal structure of the proposed framework. Individual bi-group NN have been generated for the specific regions required.

Since the classes of IMI or AMI and LVH are initially considered to be mutually exclusive, more than one mutually exclusive like bi-group NN may be asserted under one of three possible conditions: (a) the patient exhibits combined infarction; (b) the patient exhibits a combination of infarction and hypertrophy; or (c) an error has occurred. Fig. 7 represents the flowchart of operation as employed by the combination module.

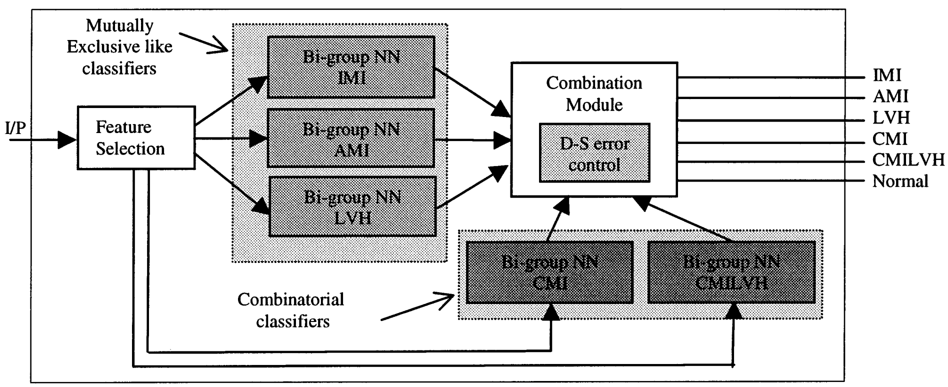

Fig. 6. Proposed classification framework employing bi-group NN and a post processing combination module.

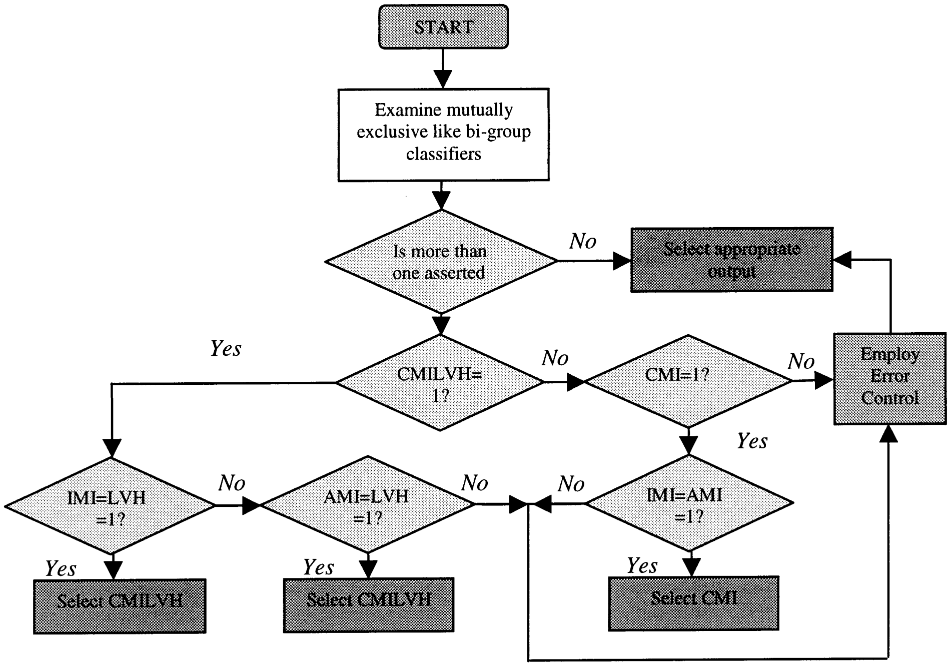

Fig. 7. Flowchart of operation of the combination module.

The combination module, upon detecting such a conflicting situation examines the outputs from the existing two combinatorial classifiers. If either one is asserted and the other appropriate mutually exclusive like networks are asserted, the appropriate overlapping class is selected. Otherwise, if no combinatorial networks are asserted, the combination module selects the error control module employing Dempster–Shafer Orthogonal Summation, hence producing one single end result.

3.3. Error control employing evidential reasoning

The error control is employed once the combination module has detected the presence of a conflicting situation i.e. two of the mutually exclusive like networks have asserted outputs yet both the remaining combinatorial classifiers have suggested no positive outputs. The error control employs a method of Evidential Reasoning based on combination performed by Dempster–Shafer Orthogonal Summation [19]. This technique establishes a framework for reasoning under uncertainty. The framework provides a facility for the combination of a number of disjoint, independent hypotheses, producing a final result based on all the available evidence which is considered to be the most likely. Considering two independent pieces of evidence $m_1(A_i)$ and $m_2(A_j)$, a combined hypothesis $m_3(A_h)$ may be produced by orthogonal summation of the form:

$$m_3(A_h) = \frac{1}{1-k} \sum_{A_i \cap A_j = A_h} m_1(A_i) m_2(A_j)$$

$$k = \sum_{A_i \cap A_j = 0} m_1(A_i) m_2(A_j)$$

where m_x is considered to be the output from each conflicting network, representing the amount of support given to the possibility of diagnostic class A_y. The framework facilitates the adjustment of pieces of evidence, supporting assessment of the credibility of the evidence source, prior to summation. This adjustment is referred to as discounting, whereby a new evidence source is produced by weighting the original value by some constant $\alpha \in R[0,1]$,

$$m^{\text{discounted}}(A_j) = \alpha m(A_j)$$

Thus, prior to orthogonal summation, by employing discounting it is possible to assess the credibility of each individual conflicting bi-group classifier. To avoid a heuristic approach to discounting factor selection, a proportion of the entire database was partitioned to allow for generation of discounting factors. Fig. 8 indicates the segregation of the database.

Following training of the networks, they were subsequently tested on unseen data. The discounting factor α_i, for each individual network was derived from the percentage of records correctly classified by the networks following exposure to the unseen data.

$$\alpha_i = \frac{t_p + t_n}{n} \times 100\%$$

where: t_p, true positive; t_n, true negative; n, number of records; i, diagnostic class.

It has been shown [15] that a positive moderation to the discounting factor α_i will positively enhance the effect of m_i for class C_i following orthogonal summation. Thus the larger value of m_i, the larger contribution it will have to the overall final result. The effect of increasing α_i will subsequently reduce the contribution of m_j, where m_i and m_j are the two conflicting outputs and the required class is C_i. This procedure hence facilitates the conflicting situation.

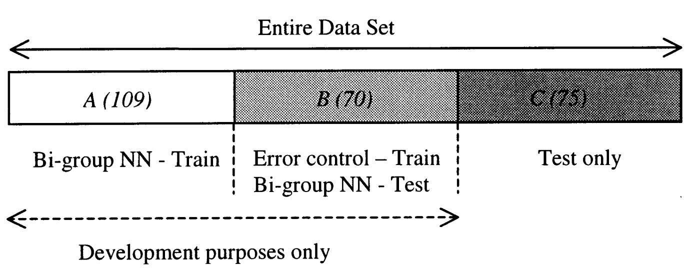

Fig. 8. Segregation of entire data into a training set for the NN, a training set for the Evidential Reasoning framework and a final test set.

4. Methods and results

The individual mutually exclusive bi-group NN were trained with individual subsets from 109 records. Where possible, equal distributions of the individual diagnostic classes were maintained in the subsets. The networks were then tested on 70 records, obtaining the appropriate discounting factors. The remaining two combinatorial networks, employed to assist with conflicting situations were trained on subsets of data from the train and initial test sets of data. This was employed to try and maximise their generalisation abilities on unseen data since it was not required to generate discounting factors for these networks. The entire system was tested on a set of 75 records. Standard Back Propagation was employed as the training algorithm, used to train two layer, single output multi-layer perceptrons. Each of the networks were trained ten times, five times with different hidden layer architectures and another five times with the same varying internal configurations but with different random start conditions. For the mutually exclusive like networks, five different methods of selecting the appropriate feature vector were investigated.

4.1. Feature selection techniques

The aim of the feature selection was to determine suitable input feature vectors which would discriminate between the required bi-group diagnostic classes. The adverse effects of this technique results in the generation of a smaller descriptive feature vector and hence subsequently reducing the architecture of the network itself and increasing its generalisation ability. The following feature selection techniques were investigated.

(1) Selection of criteria based on conventional electrocardiograph diagnostic heuristic rules. This method involved the selection of the features, as would be considered by a human expert. Since this is the standard approach, it was used as a benchmark for all other techniques, investigating if it was possible to reduce the original number of features without loss of classification accuracy.
(2) Forward Stepwise Multiple Linear Regression. This technique selects the most statistically significant features based on those indicated by the regression analysis which are to be included in the final linear regression equation.
(3) Prior to performing the regression in (2), the training portion of the database was statistically screened, based on the procedures described by Tabachnick et al. [21] removing any information, or records which may be considered to bias the regression analysis.
(4) Linear segregation based on Box Plots. This novel method involved feature selection based on linear segregation of bi-group populations between the 25th and 75th percentile ranges. If the two ranges were clearly separable, the feature retained its position in the vector, otherwise it was removed. Fig. 9 displays the two possible scenarios.
(5) For the techniques above the inputs to the networks were normalised. An additional technique was considered whereby a non-linear multiplication was

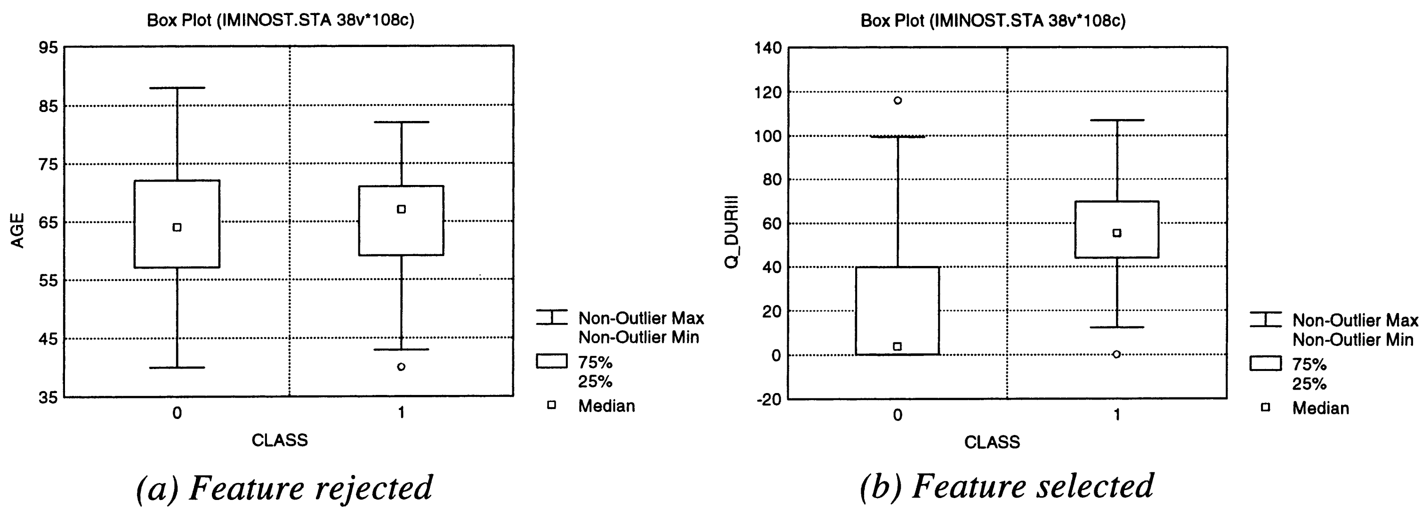

Fig. 9. Analysis of features employing Box Plots. (a) Indicates no clear discrimination between the 25th and 75th percentile ranges of the bi-group populations. (b) However, indicates a clear segregation, hence the feature is selected.

Table 1
The results for the various feature selection techniques, for the mutually exclusive like classifiers

Selection Type	IMI (%)	AMI (%)	LVH (%)
Rule based	87.14	88.57	88.57
Statistical	90.00	90.00	88.57
Statistical (screened)	90.00	92.86	92.86
Box plot	91.43	94.29	92.86
Multiplicative I/P	90	95.71	92.86

performed prior to submission of the features to the network's input. The coefficients for multiplication were obtained from the coefficients of the linear regression equation.

4.2. Results from bi-group NN

Table 1 indicates the results for the various feature selection techniques, for the mutually exclusive like classifiers. As can bee seen from the results, in all cases the feature selection techniques enhanced the performance of the individual classifiers, with reference to the heuristic technique and in no instance was performance reduced due to the input feature vector size reduction. An interesting point to note was that the method of Box Plot feature selection produced comparable results with the statistical techniques and in two instances, provided the best classification results. The combinatorial networks were generated based on a feature selection method of linear box plot segregation producing correct classification results of 89.33% for CMI and 86.67% for CMILVH. The effects of statistical screening of the training database, removing information which from a statistical sense may be considered to bias the model, was noticeable with regard to the NN classifiers. It is accepted practice that this screening should be employed prior to generation of any classifier, which adapts itself to annotated data however in many instances it is overlooked.

The potential advantage of employing such feature selection techniques when developing NN classifiers is the reduction in the size and hence computational requirements of the network configurations. Fig. 10 indicates comparisons of network sizes for the mutually exclusive like networks for the varying feature selection techniques. As can be seen, the network size of the method employing the heuristic criteria is much greater than any of the other methods, yet it exhibits in all instances the worst performance. From this it may be reasonably deduced that the original input feature vector does contain unnecessary discriminatory information, (in terms of NN classification ability) which hinders the generalisation of the classification model.

It was found that the best generalisation of the networks following exposure to the unseen data, was obtained following a low number of training epochs. Consistently, after a generalisation optimum, when the network had reached a near

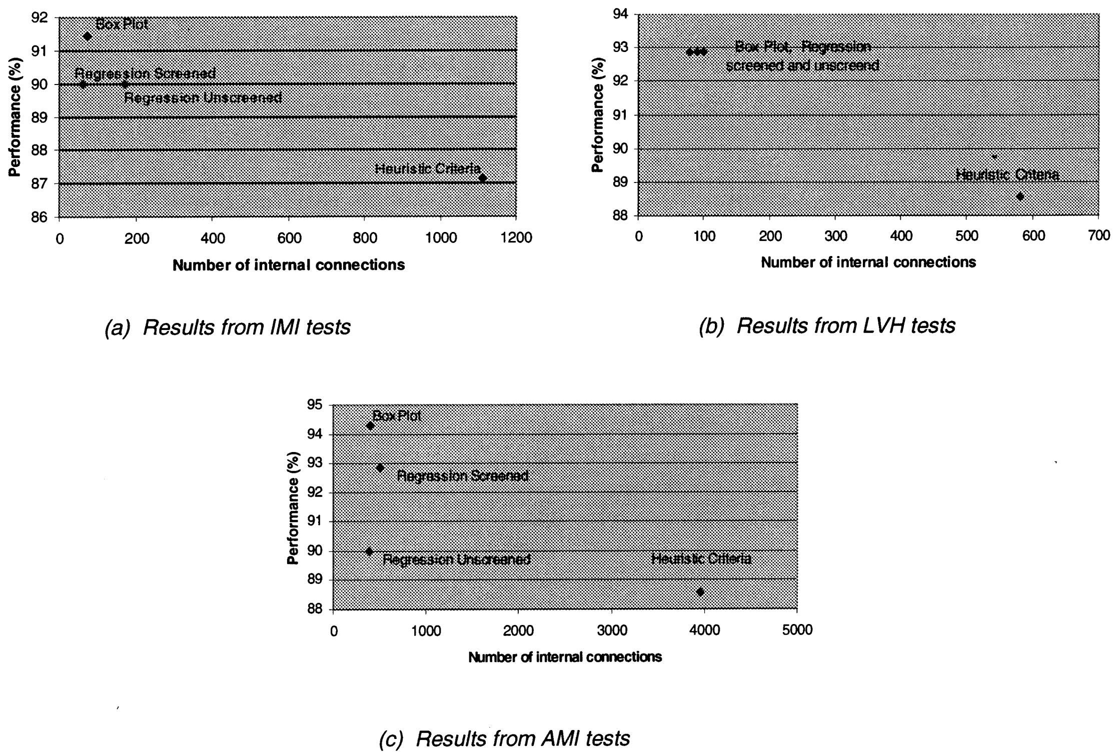

Fig. 10. Comparison of size of network and feature selection technique, with respect to classifier performance.

Table 2
Sensitivity and specificity results from framework tested on final unseen data set

	IMI (%)	AMI (%)	LVH (%)	CMI (%)	CMILVH (%)	NORM (%)
Sensitivity	66.7	75	77.8	71.4	71.4	96
Specificity	95	98.4	95.5	98.5	100	86

optimal level of performance with respect to the training data, the generalisation with regard to the test data either remained constant or decreased, thus entering into an over-train situation. Therefore, in order to obtain the optimal final trained network, rather than adopting the conventional approach of training the network until a predefined error threshold is reached, the networks were monitored over a number of training cycles and the appropriate network subsequently selected. The records employed during the Evidential Reasoning training were used as the test set. Fig. 11 is indicative of these findings, indicating regions of overtraining and location of generalisation optimums following low numbers of training epochs.

The entire framework was tested on the totally unseen final set of test records. From the validation set, discounting factors of 0.91 for IMI, 0.95 for AMI and 0.92 for LVH were assigned. An overall correct classification performance of 80% was attained. Table 2 indicates the sensitivity and specificity obtained for the individual classes. It was found from testing that the combinatorial networks provided lower classification results in comparison with the other bi-group classifiers. In an effort to monitor the problem and the adverse effects on the entire classification framework, the networks were replaced by artificial models simulating the correct desired response. With their inclusion it was found that the performance was increased, indicating good cause to include them in the framework. The error control module was replaced with a Majority Rule voting strategy, in effect removing the ability of the combination module to accommodate conflict between mutually exclusive like classifiers. The framework was also compared with a decision tree, based on a heuristic rule based set and a conventional method of NN classification, employing a multi-layer perceptron with six outputs. Results are indicated in Table 3.

As can be seen from Table 3, the proposed framework in its entirety provided significantly superior classification results compared with a conventional approach

Table 3
Results of classification frameworks on final test set

Classification technique	Performance (%)
Bi-group Framework	80.00
Bi-group Framework with artificial combinatorial networks	82.67
Bi-group Framework with no error control	78.67
Bi-group with no combinatorial networks	78.67
Heuristic Classification Framework	66.67
Conventional Multi output NN	68.00

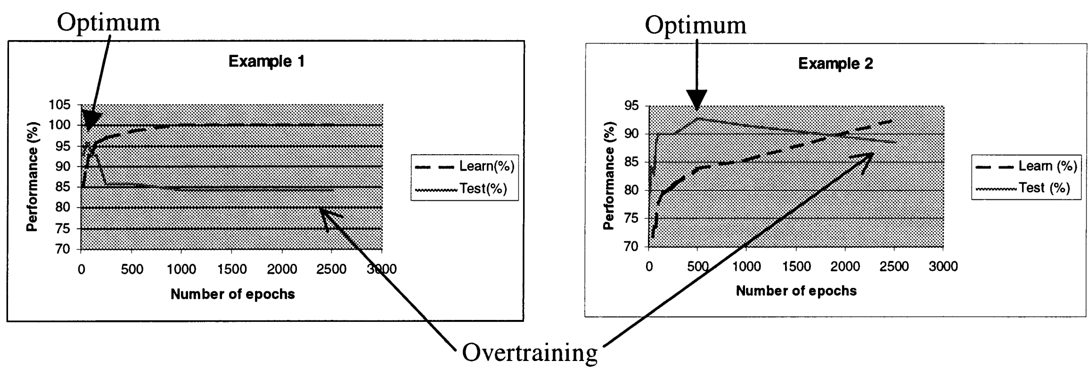

Fig. 11. Location of optimal generalisation of bi-group classifiers.

employing a multi output NN and a conventional rule based approach. It was also found that when the error control module was removed the overall classification performance was decreased, as also was the case when the combinatorial networks were not considered. The framework with the artificial combinatorial networks was seen to provide the best performance, subsequently identifying the need to further improve their generalisation behaviour.

5. Conclusions and further work

The proposed approach of the generation of a classification framework, based on sub-dividing the classification space provided significantly superior results in comparison with conventional existing techniques. The framework in its original entirety attained a correct classification level of 80% when exposed to the test data, compared with correct classification levels of 66.7% for a rule based approach and 68% for a conventional multi output NN approach. The ability to segregate the problem into a number of smaller bi-dimensional ones permitted specific attention to be paid to the selection of the most appropriate feature vector for each individual diagnostic class, subsequently increasing the generalisation of each bi-group NN classifier and additionally, decreasing the computational requirements of the network. An interesting result from the feature selection tests indicated the ability of the box plot feature selection to provide comparable results with statistical techniques. The results also indicated the need to further improve on the generalisation ability of the combinatorial networks. By removing the developed networks and replacing them with artificial correct inputs increased the classification ability of the framework to 82.67%. At this point, it may be suggested that the cause of this problem be attributed to low numbers of CMI and CMILVH classes in the training data sets in comparison with the proportion of the remaining classes. Preliminary work has also been undertaken to expand the size of the these classes in the database. It was also found that the combined evidence provided from the mutually exclusive like classifiers and the combinatorial classifiers, in comparison with a framework without the combinatorial classifiers, produced a higher level of correct classification.

The bi-group classifiers were found to reach an optimal generalisation level on the unseen data following a low number of training epochs. Work is currently being investigated to generate a prediction equation to locate the optimal number of training epochs for each individual classifier.

References

[1] Attikiouzel Y, de Silva CJS. Applications of neural networks in medicine. Aust Physical Eng Sci Med 1995;18:158–64.

[2] Bounds DG, Lloyd PJ. A comparison of neural network and other pattern recognition approaches to the diagnosis of low back disorders. Neural Networks 1990;3:583–91.

[3] Clark JW. The origin of biopotentials. In: Webster JG, editor. Medical Instrumentation Application and Design, 2. New York: Wiley, 1995.
[4] Cox JR, Nolle FM, Arthur RM. Digital analysis of the electroencephalogram, the blood pressure wave and the electrocardiogram. Proc IEEE 1972;60:1137–64.
[5] Dorffner G, Porenta G. On using feedforward neural networks for clinical diagnostic tasks. Artif Intell Med 1994;6:417–35.
[6] Egmont-Petersen M, Talmon JL, Brender J, McNair P. On the quality of neural net classifiers. Artif Intell Med 1994;6:359–81.
[7] Haykin S. Neural Networks a comprehensive foundation. New York: IEEE Press, 1994.
[8] Kilpatrick D, Johnston PR. Origin of the electrocardiogram. IEEE Eng Med Biol Mag 1994:479–486.
[9] Kors JA, van Bemmel JH. Classification methods for computerised interpretation of the electrocardiogram. Meth Inform Med 1990;29:330–6.
[10] Lim CP, Harrison RF, Kennedy RL. Applications of autonomous neural network systems to medical pattern classification tasks. Artif Intell Med 1997;11:215–39.
[11] Lippmann R. An introduction to computing with neural nets. IEEE ASAP Mag 1987;4:4–22.
[12] Lippmann R. Pattern classification using neural networks. IEEE Commun Mag 1989;27:47–64.
[13] Nugent CD, Webb JAC, Wright GTH, Black ND. Electrocardiogram 1: pre-processing prior to classification. Automedica 1998;16:263–82.
[14] Nugent CD, Webb JAC, Wright GTH, Black ND. ECG classification: a hybrid approach. Proc ISSC '97 1997:259–266.
[15] Nugent CD, Webb JAC, McIntyre M, Black ND, Wright GTH. Computerised electrocardiology employing bi-group neural networks. Artif Intell Med 1998;13:167–80.
[16] Pipberger HV. Computer analysis of the electrocardiogram. Comput Biomed Res 1965;1:377–407.
[17] Pipberger HV, McCaughin D, Littmann D, Pipberger HA, Cornfield J, Dunn RA, Batchlor CD, Berson AS. Clinical application of a second generation electrocardiographic computer program. Am J Cardiol 1975;35:597–608.
[18] Plonsey R. The biophysical basis for electrocardiography. CRC Crit Rev Bioeng 1976;1:1–48.
[19] Shafer G. A Mathematical Theory of Evidence. Princeton, NJ: Princeton University Press, 1976.
[20] Stallman FW, Pipberger HV. Automatic recognition of electrocardiographic waves by digital computer. Circ Res 1961;9:138–43.
[21] Tabachnick BG, Fidell LS. Using Multivariate Statistics. New York: Harper Collins, 1996.
[22] Widrow B, Lehr MA. Thirty years of adaptive neural networks: perceptron, madaline and backpropagation. Proc IEEE 1990;78:1415–42.
[23] Willems JL, et al. Effect of combining electrocardiographic interpretation results on diagnostic accuracy. Eur Heart J 1988;9:1348–55.
[24] Willems JL, Arnaud P, van Bemmel JH, Degani R, MacFarlane PW, Zywietz C. Common standards for quantitative electrocardiography: goals and gain results. Meth Inf Med 1990;29:263–71.
[25] van Bemmel JH, Zywietz C, Kors JA. Signal analysis for ecg interpretation. Meth Inf Med 1990;29:317–29.