Detection of Ventricular Ectopic Beats Using Neural Networks

Hann-Shi Chow, George B. Moody, and Roger G. Mark
Harvard-MIT Division of Health Sciences and Technology, Cambridge, Massachusetts USA

Abstract

We describe a system of three artificial neural networks (ANNs) trained to detect ventricular ectopic beats (VEBs) in the AHA Database. Two ANNs are trained for each record (one for each input signal), using a learning set containing five normal beats obtained from the record and seven VEBs obtained from a larger database. These two ANNs function as pattern recognizers for each record. The third ANN is trained to arbitrate disagreements between the first two ANNs. We replaced the beat classification logic of an existing VEB detector with this system of ANNs and evaluated its performance using 69 records from the AHA Database (only records with paced rhythm or ventricular fibrillation were excluded). Gross VEB sensitivity improved from 94.83% to 97.39% when the ANN system replaced the original beat classification logic, and gross VEB positive predictivity improved from 92.70% to 93.58%.

1 Introduction

During the past few years, artificial neural networks have been widely studied in the hope of achieving human levels of performance in applications such as pattern recognition, image processing, robotics, and knowledge extraction, where conventional algorithms generally do not approach such levels.

Pattern classifiers exist in many forms. The task of any pattern classifier is to arrange the input data into subsets. ANNs represent an important subset of these classifiers. ANNs are attractive for solving pattern recognition problems because they require very few assumptions about the underlying data. The development of a successful ANN-based pattern classifier requires a well-constructed, accurately characterized, and representative set of patterns for training; pre-processing and post-processing algorithms; an appropriate network topology; and an evaluation database [1].

In analysis of the electrocardiogram, it is widely accepted that a multi-lead analysis has substantial advantages over a single-lead analysis [2]. First, although noise cannot be avoided entirely, it is less likely to occur simultaneously on multiple independent leads than on any single lead [3]. Second, although it is possible for different types of QRS complexes to appear similar in a single lead, this situation is much less likely to occur simultaneously in multiple independent leads. Third, although a QRS complex may be nearly isoelectric, hence difficult to detect, in a single lead, this is unlikely to be the case in multiple independent leads.

We describe here the design and evaluation of a technique for detecting VEBs from two ECG leads using a system of three back-propagation ANNs [4]. (In principle, this technique is extensible to n leads using $n+1$ ANNs; here we discuss only the two-lead case.) This set of three ANNs replaces the pattern recognition and the beat classifier parts of a conventional algorithm developed by the first author. We compare the performance of this modified algorithm (ANN3) with a simpler version using only two ANNs (ANN2), with the conventional algorithm (HSC), and with another algorithm developed by the second author (ARISTOTLE). All evaluations were performed using records 1201-7210 of the AHA Database for Evaluation of Ventricular Arrhythmia Detectors [5] (69 records; records 2202 and 8201-8210 were excluded because of the presence of paced rhythm or ventricular fibrillation). All performance statistics reported here are "gross" statistics obtained by the methods specified by the AAMI [6].

2 Methods

VEB detection algorithms are usually composed of three distinct parts: QRS detection, pattern recogni-

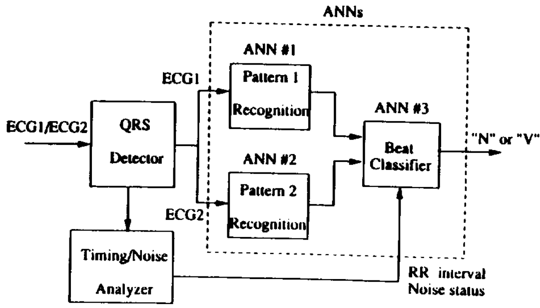

Figure 1: Structure of the VEB detection algorithm ANN3.

Figure 2: Seven VEBs used to train all ANNs.

tion (morphology clustering), and beat classification. Figure 1 shows the structure of algorithm ANN3 in these terms.

2.1 Using ANNs for Pattern Recognition

Each ANN used in this study consists of 16 input perceptrons, 6 hidden-layer neurons, and a single output neuron. Each detected beat supplies a set of inputs. For each ANN, the 16 inputs are normalized samples from a single ECG signal, taken at 8 ms intervals beginning 48 ms before the detected R-wave peak. Some form of normalization is necessary in any case to match the input signals to the perceptron input range; a well-chosen normalization method reduces the deleterious effects of baseline shifts and variations in signal amplitude on algorithm performance. We normalize each set of inputs such that its minimum value is zero and its maximum is one. Two ANNs are trained for each ECG record (one for each lead), using a learning set containing five normal beats and seven VEBs.

The seven VEBs are the same in each case, having been chosen from the AHA DB to represent common VEB morphologies (Figure 2). A new set of five normal beats is chosen by the algorithm for each record, using an automatic learning process based on measurements of RR intervals and QRS duration. Once the dominant normal beat type has been identified by this process, a clean example is selected from the record, and four others are synthesized from it by adding small amounts of noise to it. ANN training then begins, using the seven VEBs and the five normal beats as the learning set. During training, beats in the learning set are presented as ANN inputs, and back-propagation is used to adjust the network weights until the ANN produces an output near zero for a normal input and near one for a VEB input. Based on the actual values obtained during training, a threshold for beat classification is set.

Once the ANNs have been trained, they are used to classify the remainder of the beats. For each beat, the output of each ANN is a value between zero and one. These values are compared to threshold values, in order to obtain tentative beat classifications. If the tentative classifications agree, they are accepted without further analysis. Otherwise, some means of arbitration must be used to obtain a final classification.

2.2 Resolving Discrepancies

The problem of arbitrating between discrepant decisions is often quite difficult. For algorithms HSC and ANN2, we employ a simple fixed strategy (a beat called a VEB by only one ANN is labelled normal unless it is premature and followed by a compensatory pause). This simple strategy works well when applied to the ANN outputs, but we wanted to explore more general approaches.

For algorithm ANN3, therefore, we used a third ANN to perform the arbitration (or "fusion") function. This ANN consists of eight input perceptrons, six hidden-layer neurons, and a single output neuron. Four kinds of inputs are derived from the signal: (1) the output value from the pattern-recognition ANN; (2) the measures of beat timing including premature and compensatory pause; (3) a function related to the amount of baseline wander; and (4) a function related to the amount of high frequency noise.

The amount of baseline wander for beat j is estimated by:

$$N_{bw}[j] = |bw[j-1] - bw[j]| + |bw[j+1] - bw[j]| \quad (1)$$

where $bw[j]$ is the average of 8 successive samples at 4 ms intervals during an interval centered between the

R-wave peaks of beat $j-1$ and beat j. High frequency noise is estimated from 32 successive samples centered on the same time [2]:

$$N_{em} = \frac{1}{32}\sum_{i=1}^{n} |x[i] - x_{pred}[i]| \tag{2}$$

where $x[i], i = 1, \ldots, 32$ are the input samples, and $x_{pred}[i]$ is predicted by linear interpolation between $x[1]$ and $x[n]$. Before supplying these estimates to the third ANN, they are normalized by an estimate of the peak-to-peak amplitude of the dominant normal beats.

This network requires its own learning set; in this case we defined a fixed learning set, since the target outputs needed for training cannot be determined by an automated procedure, as is possible for the pattern-recognition ANNs. The learning set consists of beats (both VEBs and non-VEBs) for which tentative classifications differ. In order to obtain sufficient numbers of such beats, and a better representation of noise, we used the noise stress test method [7] to add baseline wander and electrode motion artifact to selected records in the AHA DB. From a total of 77 records (69 original records as above, and 8 to which noise was added), we extracted all discrepancies from the first five minutes of reference-annotated ECGs (ignoring discrepancies that occurred during the segments at the beginning of each record, for which no reference annotations are available). These beats (567 VEBs and 1313 other beats) were used to train the third ANN.

Detector	VEB Se	VEB +P
ARISTOTLE	98.65%	94.22%
HSC	94.83%	92.72%
ANN2	96.89%	92.35%
ANN3	97.39%	93.58%

Table 1: Summary of VEB detector performance on 69 AHA DB records.

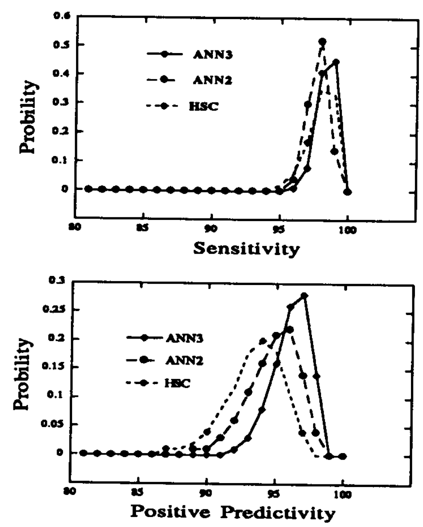

Figure 3: Bootstrap distributions of gross VEB sensitivity (above) and gross VEB positive predictivity (below).

3 Results

3.1 AHA Database tests

Using standard software for evaluating arrhythmia detectors [8], we compared the gross VEB sensitivity (Se) and positive predictivity (+P) of algorithms HSC, ANN2, and ANN3 on the 69-record subset of the AHA DB (Table 1). These three algorithms all use the same QRS detector (on these records, the gross QRS Se is 99.86%, and the gross QRS +P is 99.63%). Differences in VEB detection are attributable to the use of neural networks by algorithms ANN2 and ANN3. As a further point of reference, Table 1 also includes statistics for ARISTOTLE on the same 69 records; in this case, differences are attributable in part to a different QRS detector (for which the gross QRS Se is 99.94% and the gross QRS +P is 99.86% on these records). In all cases, both leads were analyzed.

For algorithms HSC, ANN2, and ANN3, we also derived bootstrap estimates of the distributions of these statistics [9]. Figure 3 shows the estimated distributions of VEB sensitivity and VEB positive predictivity, based on 10,000 bootstrap iterations. It is apparent that ANN3 performs significantly better in these tests than HSC and ANN2; the difference between HSC and ANN2 does not appear to be significant in these tests.

3.2 Noise stress tests

It is useful to characterize the noise level in a noise stress test in terms of the signal-to-noise ratio (SNR) during the noisy segments. SNR is commonly expressed in decibels (dB):

$$\mathrm{SNR} = 10 \log(S/N) \tag{3}$$

where S is the power of the signal, and N is the power of the noise. In the noise stress test, S is defined as a trimmed mean of the normal QRS peak-to-peak amplitude, and a frequency-weighted RMS noise estimate is used for N [8].

In a series of tests, we added baseline wander and electrode motion artifact to each signal alternately,

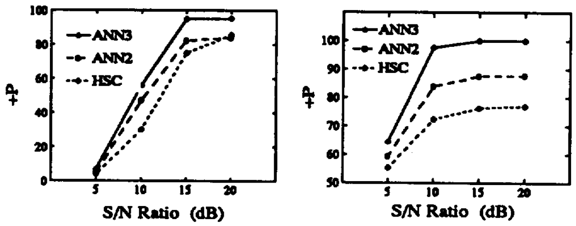

Figure 4: Dependence of VEB positive predictivity (VEB +P) on signal-to-noise ratio (SNR), in noise stress tests using AHA DB record 7204 with added electrode motion artifact (left) and baseline wander (right).

varying the SNR between tests. Figure 4 summarizes the dependence of VEB positive predictivity upon SNR for algorithms HSC, ANN2, and ANN3, in two sets of noise stress tests for which we used AHA record 7204. (VEB positive predictivity, related to false positive VEB rate, is a good index of a VEB detector's noise tolerance, since the most common errors in noise are likely to be false positive VEBs.) The figure shows the marked superiority of algorithm ANN3, which maintains an excellent positive predictivity even when the signals are quite noisy. Algorithm ANN2 tolerates noise better than HSC, but substantially worse than ANN3.

4 Discussion

The ANN3 algorithm, incorporating automated algorithms for patient-specific ANN training, performs well. On average, only 13 seconds of signals were required for training, so ANN3 may be useful for real-time analysis. The computational requirements for training are likely to be much more than 13 seconds using the resources of typical single-patient arrhythmia monitors, but the amount of computation required once training is complete is comparable to that needed by conventional algorithms. This suggests the possibility of a hybrid approach in which the training is performed remotely on hardware optimized for this task, after which the network weights can be downloaded to a small real-time processor.

In this study, only seven typical VEBs were used for training. It is possible that better performance might be obtained by increasing the size of the learning set, but doing so would lengthen training times.

From these results it is apparent that the use of neural networks by ANN3, and to a lesser extent, ANN2, provides a performance advantage over the conventional HSC algorithm. At least some of the difference between the performance of ANN3 and that of ARISTOTLE may be attributable to ARISTOTLE's QRS detector. It should be noted that both HSC and ARISTOTLE have been the results of many years of research; the present study produced results we consider remarkable given the relatively modest (though non-trivial) effort required to obtain them. This observation suggests that, while conventional algorithms such as ARISTOTLE may still outperform ANN-based methods, ANNs offer a method of approaching such levels of performance at low cost.

5 Acknowledgement

This work was supported by Nihon Kohden Corporation, in Tokyo, Japan.

References

[1] S.G. Artis, R.G. Mark, and G.B Moody. Detection of atrial fibrillation using artificial neural networks. *Computers in Cardiology*, pages 173–176, 1991.

[2] G.B Moody and R.G Mark. Development and evaluation of a 2-lead ECG analysis program. *Computers in Cardiology*, pages 39–44, 1982.

[3] R.M. Armington and F.J. Geheb. Optimizing the utilization of multi-lead ECG data for accurate QRS detection. *Computers in Cardiology*, pages 259–262, 1987.

[4] R.Y. Levine and T.S. Khuon. *Decision-level neural net sensor fusion*, volume 10. North Hollord Pub. Co., 1992. Handbook of statistics.

[5] AHA Database for Evaluation of Ventricular Arrhythmia Detectors. Available from: ECRI, 5200 Butler Pike, Plymouth Meeting, PA 19462 USA.

[6] AAMI. *Testing and Reporting Performance Results of Ventricular Arrhythmia Detection Algorithms*. Association for the Advancement of Medical Instrumentation, Arlington, VA, 1987. Publication ECAR-1987.

[7] G.B. Moody, W.K. Muldrow, and R.G. Mark. A noise stress test for arrhythmia detectors. *Computers in Cardiology*, pages 381–384, 1984.

[8] DB Software Package. Available from: MIT-BIH Database Distribution, MIT Room 20A-113, Cambridge, MA 02139 USA.

[9] P. Albrecht, G.B. Moody, and R.G. Mark. Use of the 'bootstrap' to assess the robustness of the performance statistics of an arrhythmia detector. *J. Ambulatory Monitoring*, 1(2):171–176, 1988.