Automated ECG Analysis using Weightless Neural Networks

Wouter A.Th. Manintveld[1,2], *Igor Aleksander*[2], *Léon J.M. Rothkrantz*[1]

[1] *Delft University of Technology, Department of Technical Mathematics and Informatics, Zuidplantsoen 4, 2628 BZ, Delft, The Netherlands, alparon@kgs.twi.tudelft.nl*
[2] *Imperial College of Science, Technology and Medicine, Department of Electrical and Electronic Engineering, Exhibition Road, SW7 2AZ, London, United Kingdom*

ABSTRACT

A Weightless Neural Network has been used for the classification of electrocardiograms (ECGs). The algorithm was tested on the standard MIT-BIH database. Weightless Neural Networks are based on the n-tuple classification method and are very powerful tools for pattern recognition tasks, owing to their speed, for both training and testing, and their training capabilities.

The method used consists of two stages: preprocessing of the data, and a neural network based classification stage. For each class of heartbeats a weightless discriminator is trained, and during operation the discriminator with the highest output determines the class of the input pattern.

The results of a one-lead ECG classification system, based on this rather brute force approach, were encouraging. With a correct classification of approximately 97 percent, this system can be a suitable basis for an ECG-diagnosis system.

1 Introduction

Heart diseases are the most common cause of death in the western world. Patients with heart related diseases need considerable medical attention, because of the constant danger of complete heart failure. Electrocardiography is an effective non-invasive diagnostic technique, which is widely used by cardiologists. Electrocardiograms are electrical views of the heart, recorded by placing electrodes on the patient's body. ECGs provide the cardiologist with useful information about the rhythm and functioning of the heart.

In recent years a lot of research has been done to assist cardiologists with their task of diagnosing ECG recordings. The research fields range from novelty detectors to fully automated ECG diagnosing systems. A wide range of techniques has been used, including statistical pattern recognition, expert systems and artificial neural networks.

A very powerful, but somewhat unorthodox neural technique for pattern recognition is the Weightless Neural Network. Weightless Neural Networks are called 'unorthodox' because of the omission of weights. Instead a look-up table approach is used.

In this project the use of Weightless Neural Networks in ECG analysis was examined. The goal was to classify beats extracted from ECG signals into several predefined classes, based on the morphology of the data.

2 Weightless Neural Networks

Weightless Neural Networks are based on a different approach than the 'conventional' neural networks. The basic idea is that the network is trained by storing the training samples in memory, rather than adjusting weights in order to fit the samples better. Weightless Neural Networks originate from the n-tuple sampling machine introduced by Bledsoe and Browning in 1959 [3]. This n-tuple sampling machine was used for character recognition, using several n-bit vectors as input for their system. In such a n-bit vector a bit was set if a corresponding photo-cell detected a black spot on the input pattern.

2.1 The RAM-node

The RAM-node [1, 2] is the basic component of a Weightless Neural Network. It is, as usual for a neural building brick, a very simple element (see figure 1).

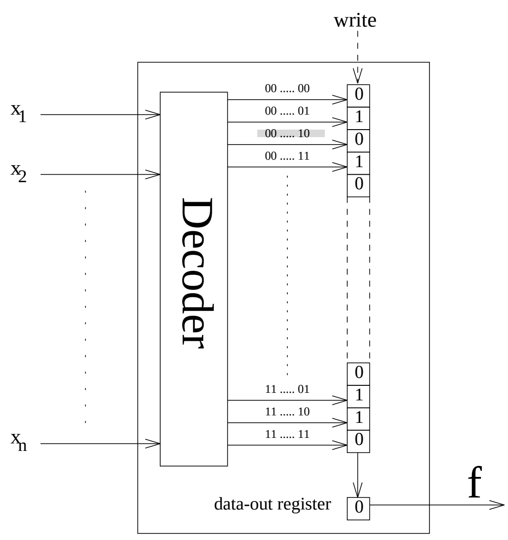

Fig. 1: A RAM-node

Its function is to determine if a bit vector is present in the set of bit vectors it is trained with. In order to do so it uses a so called look-up table. This look-up table can easily be implemented in the conventional memory of a computer, which makes learning a one-shot memorisation task.

A RAM-node has no real generalisation: it has a binary output which merely tells if a pattern has occurred in the training data. Although this seems a disadvantage, it can easily be solved by combining several RAM-nodes.

2.2 Discriminator

A discriminator consists of a layer of RAM-nodes organised as shown in figure 2. The most important feature of a discriminator is that it splits the input pattern into K n-sized tuples, each of which are used as the input for a RAM-node. In this way the input pattern of the discriminator is distributed over the memory of several RAM-nodes, instead of being stored in only one location.

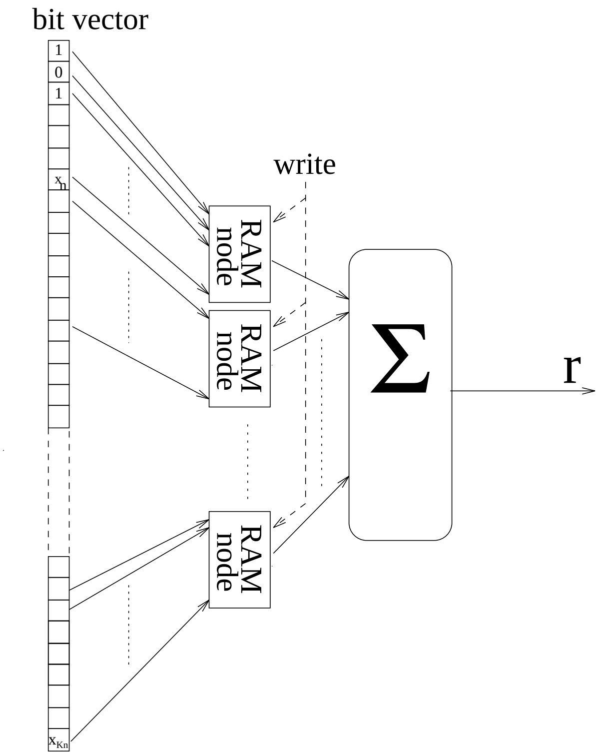

Fig. 2: A discriminator

WISARD is a binary pattern recognition system using several weightless discriminators. *WISARD* stands for *WI*lkie, *S*tonham and *A*leksander's *R*ecognition *D*evice.

The WISARD system uses a simple, supervised training algorithm. When a bit vector and its class are presented, it only trains the discriminator assigned with the task of learning patterns belonging to that particular class.

3 Data

Record 208 from the MIT-BIH database [4] was used to test the system. The patient from whom this ECG was recorded was a woman aged 23. The two leads are MLII and V1. The signal contains three different kinds of beats: Normal Beats (N), Premature Ventricular Contractions (PVC or V) and Fusion beats (F). This particular record was selected because it contains more than two classes; it has some noise periods, and due to the availability of results obtained from other classification methods.

An analysis of the raw data was made using the algorithm of John W. Sammon jr., as described in [6]. The algorithm is based upon a point mapping of N vectors from the $\mathcal{R}^n$-space to a lower-dimensional space such that the inherent structure of the data is approximately preserved under the mapping. The approximate structure preservation is maintained by fitting N points in the lower-dimensional space such that their interpoint distances approximate the corresponding distances in the $\mathcal{R}^n$-space .

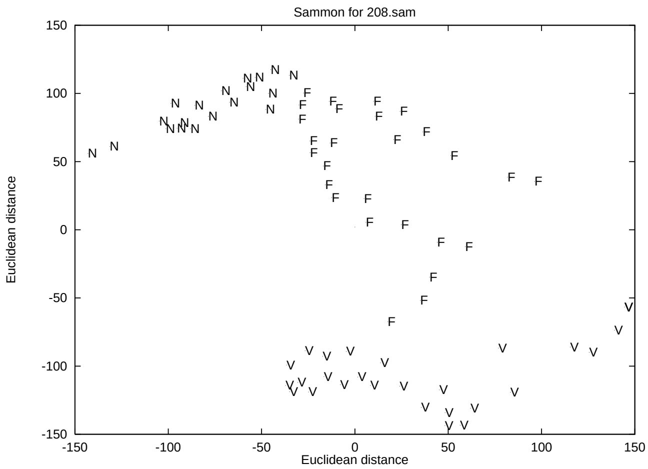

Fig. 3: A Sammon mapping of signal 208, using 100 beats of each class

The Sammon mapping of 100 beats of each class(figure 3) shows that there is some inherit clustering in the data, although the clustering seems rather loose, especially for the Fusion beats. Accordingly, classification of the beats should be possible. Neural network should be able to find some borders between the classes in the multidimensional space, using the raw data as their input.

4 System Design

The classification system under study is based on the WISARD type multi-discriminator architecture: for each relevant kind of beat a discriminator is trained. Figure 4 shows the design of the system.

The ECG data is processed and *thermometer* coded in order to feed the data into the discriminators. The system for classification of ECG beats consists of two parts: the first is the preprocessing part, the second the weightless neural network part. The main function of the preprocessing part is the coding of the data into a big string.

During the supervised training phase of the second, neural part, each discriminator learns a specific kind of beat. During the test phase the discriminator with the highest output determines the type of the beat fed into the system.

One beat is taken to be a fixed number of sample values (or, equivalently, milliseconds). The number of samples is called 'wsize', from window size. As the default value for $wsize$ 300 was taken; this corresponds to 0.833 seconds of the signal, the duration of one beat at a normal rate of 72 beats per minute[1].

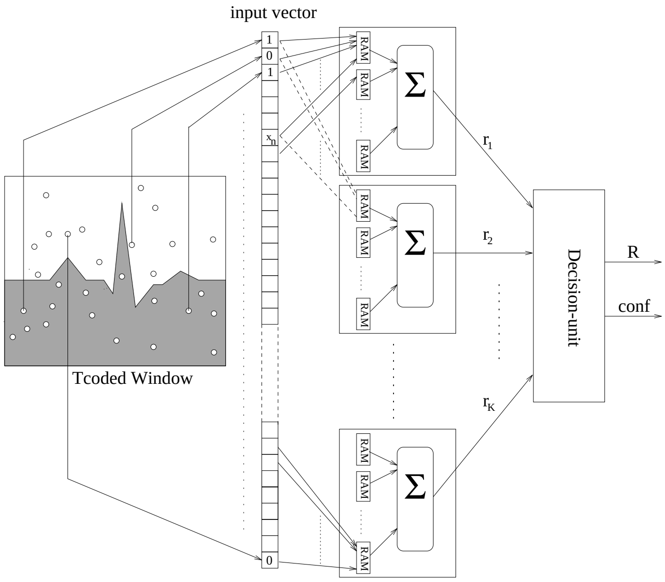

Fig. 4: The system design

A retina, an array of input points, is put over the window. Each point in the retina is related to one bit in the input vector of the neural network. If a retina point lies below the signal in the window, the corresponding bit in the vector is set to 1, otherwise the bit is reset to 0. This method is called *thermometer coding* because it resembles the way mercury points out a temperature in a thermometer: black ("ones") under, and white ("zeros") above the value.

In order to incorporate prior knowledge about the shape of the ECG beats, the retina points are selected partly uniformly random from the window, and partly normally distributed along a chosen axis (figure 5), to which the beat is fitted according:

$$x_{T'} = \left\lfloor \frac{x - T_{\min}}{T_{\max} - T_{\min}} \times wres + \frac{1}{2} \right\rfloor \tag{1}$$

$T_{\min}$ is the lower bound for the sample values and $T_{\max}$ the upper bound. This equation scales the sample values to fit the window resolution ($wres$) more closely.

[1]This is the case for the MIT-BIH ECG records: these are sampled at a frequency of 360 Hz.

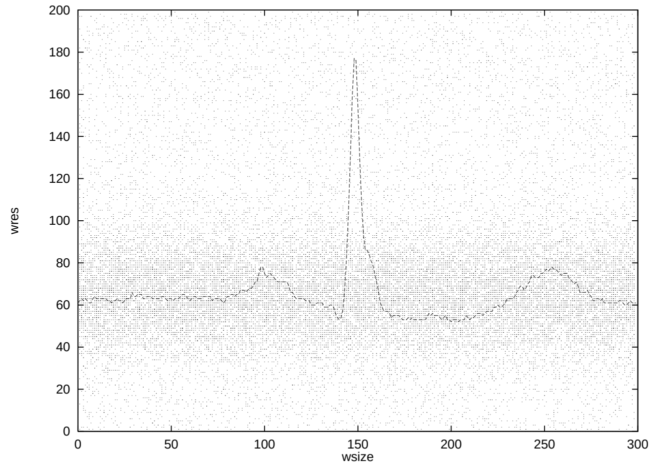

Fig. 5: A windowed beat with partly normal distributed retina

The points of the retina are randomly connected with the bits in the input vector, and covers a percentage ($coverage$) of the window points. The parameters $wsize$, $wres$ and $coverage$ determine the size of the input vector: $wsize$ and $wres$ form a grid of $wsize \times wres$ points, and $coverage$ determines the number of these points that are actually connected to a bit in the input vector.

5 Results

The network was trained with the first 5 minutes of the record 208, and tested on the remaining 25 minutes (as described in [5]). Figure 6 shows the averaged results of the experiments with the basic system for $coverage$=25, 50, 75 and 100%. The signal was scaled to a $wres$=200, using $T_{\min}$=300 and $T_{\max}$=1800.

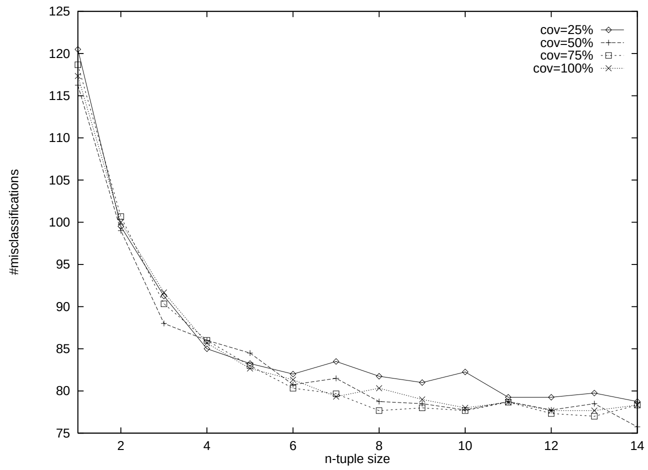

Fig. 6: Test results with 'Normal' retina

The highest classification rate seen was 97.0% for n = 14 and $coverage$=50%. The number of misclassifications was 73 out of 2437 beats. Table 1 shows the classifications of this experiment and their sensitivity and specificity. The entries represent the number of beats from

	N	PVC	F
N	1303	0	4
PVC	5	791	28
F	15	16	270
Q	3	1	1
Sens.	99.69	96.00	89.70
Spec.	98.27	97.90	89.11

TABLE 1: RESULTS

the row header class which are classified by the system as belonging to the class of the column header.

In order to obtain more insight in the way the results vary on changing the parameters, a series of experiments was performed for *coverage* ranging from 10 to 100% (stepsize 10%) and n-tuple = 1 to 16. Figure 7 illustrates that the results for *coverage* 50 – 100 % and n-tuple size 9 – 16 are approximately the same: apparently the system is insensitive to the choice of the parameters in this range.

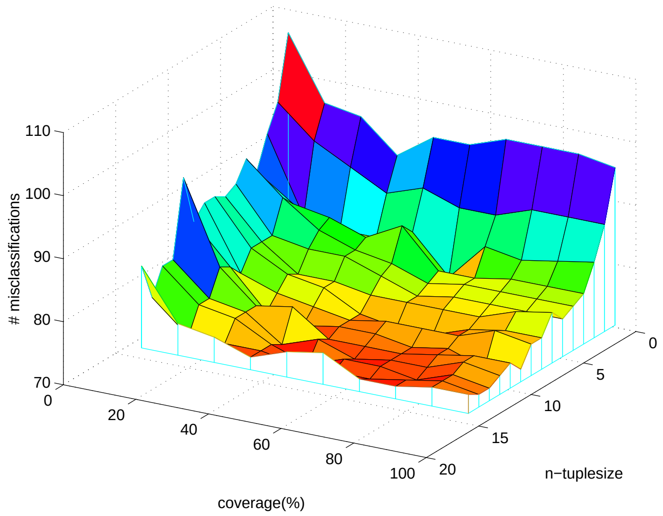

Fig. 7: Test results for *coverage* = 10–100% and n = 2–16

Figure 8 shows the average *confidences* for the previous experiments. The confidence is the relative difference between the highest and the second output of the discriminators. The confidence is averaged for all the (in)correct classifications for a certain n-tuple and *coverage*. The upper plane is the average confidence of the correct classifications and the lower the average confidence of the misclassifications.

The figure illustrates that both the confidence measures increase if the n-tuple size increases, and that the *coverage* has hardly any influence. The relation between the misclassification confidence and the correct classification confidence seems to be linear: the former is approximately half the latter. This makes the confidence an unsuitable measure for selecting the parameter values.

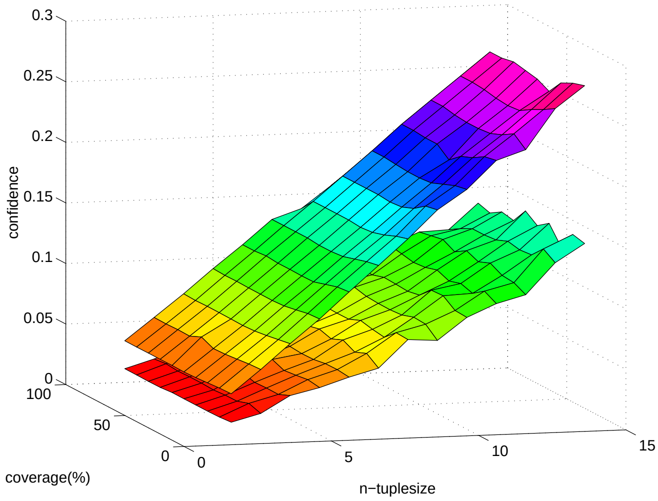

Fig. 8: Confidence planes for correct and incorrect classifications

6 Conclusions

Weightless neural techniques are fast and easily implementable neural classifiers. The results of the brute force approach used here for the classification of ECG beats are encouraging, and the insensitivity towards the parameters makes it a robust method for ECG analysis.

Applications of ECG analysis can be used in the telemedicine field. Patients can perform their own investigation of the condition of the heart by telephone. Systems like this are available, but are based on different techniques. The weightless Neural Network based approach is ideal for this task because of its robustness and their relative quick and high performance.

References

[1] I. Aleksander and H. Morton. *An Introduction to Neural Computing*. Chapman & Hall, 1990.

[2] I. Aleksander and T.J. Stonham. Guide to pattern recognition using random access memories. *IEE proceedings*, 126:42–49, 1979.

[3] W.W. Bledsoe and I. Browning. Pattern recognition and reading by machine. *Proceedings of the Eastern Joint Computer Conference*, pages 225–232, 1959.

[4] MIT-BIH database distribution, Massachusetts Institute of Technology. 77 Massachusetts Avenue, Cambridge MA 02139. USA. *MIT-BIH Arrhythmia Database Directory*, 3rd edition, 1992.

[5] George B. Moody. *Evaluating ECG Analyzers*. Harvard-MIT Division of Health Sciences and Technology.

[6] John W. Sammon jr. A nonlear mapping for data structure analysis. *IEEE Transaction on Computers*, C-18(5):401–409, May 1969.