Am. J Cardiol 1995;75:929-933
LU TP 94-14

# Artificial Neural Networks for Recognition of Electrocardiographic Lead Reversal

Bo Hedén, MD[1], Mattias Ohlsson, MSc[2], Lars Edenbrandt , MD, PhD[1], Ralf Rittner, MSc[1], Olle Pahlm, MD, PhD[1], Carsten Peterson, PhD[2]

Departments of [1]Clinical Physiology and [2]Theoretical Physics, Lund University, Lund, Sweden

Running head: Neural Networks for Analysis of Electrocardiograms

This study was supported by grants from the Swedish Medical Research Council (B94-14W-09893-03A), Swedish National Board for Industrial and Technical Development and from the Faculty of Medicine, Lund University, Sweden. The Göran Gustafsson Foundation for Research in National Science and Medicine and the Swedish Natural Science Research Council are also acknowledged for financial support.

Address for reprints
Bo Hedén
Department of Clinical Physiology
University Hospital
S-221 85 Lund
Sweden

Telephone +46 46 17 33 40
Telefax +46 46 15 17 69

## Abstract

Misplacement of electrodes during the recording of an electrocardiogram (ECG) can cause an incorrect interpretation, misdiagnosis and subsequent lack of proper treatment. The purpose of this study was twofold. (1) Develop artificial neural networks that yield peak sensitivity for the recognition of right/left arm lead reversal at a very high specificity. (2) Compare the performances of the networks to those of 2 widely used rule based interpretation programs. The study was based on 11,009 ECGs recorded on patients at an emergency department, using computerized electrocardiographs. Each of the ECGs was used to computationally generate an ECG with right/left arm lead reversal. Neural networks were trained to detect ECGs with right/left arm lead reversal. Different networks and rule based criteria were used depending on the presence or absence of P waves. The networks and the criteria all showed a very high specificity (99.87-100%). The neural networks performed better than the rule based criteria both when P waves were present (sensitivity 99.1%) or absent (sensitivity 94.5%). The corresponding sensitivities for the best criteria were 93.9% and 39.3%, respectively. An estimated 300 million ECGs are recorded annually in the world. The majority of these recordings are performed using computerized electrocardiographs, which include algorithms for detection of right/left arm lead reversals. In this study neural networks performed better than conventional algorithms and the differences in sensitivity could result in 100,000-400,000 right/left arm lead reversals being detected by networks but not by conventional interpretation programs.

## Key words

Electrocardiography

Artificial intelligence

Neural networks

Diagnosis, computer-assisted

## Introduction

Both clinicians and interpretation programs require correct data for a correct electrocardiogram (ECG) interpretation. Misplacement of electrodes during the electrocardiographic recording is one situation which results in incorrect data. Treatment errors due to lead reversal have been reported; misdiagnosis and subsequent lack of proper treatment as well as inappropriate treatment due to false positive diagnoses do occur.[1] To prevent a false interpretation, the misplacement should be recognized and corrected by the technician who records the electrocardiogram. To recognize the lead reversal from the appearance of the ECG may be difficult. Even trained ECG readers often fail to recognize a lead reversal.[2] Therefore, interpretation programs contain algorithms for detecting the most common type of lead reversal, the right/left arm lead reversal. Two widely used algorithms for the detection of right/left arm lead reversal are known to have a high specificity but rather low sensitivity. Even though it is easy for the experienced ECG reader to detect an ECG with an lead reversal, improvement of the algorithms, which are rule based, is a difficult task also for the expert. Artificial neural networks (ANNs) have proved to be very powerful in pattern recognition tasks since they can handle almost any non-linear dependence. ANNs mimic in a very crude way mammal information processing. Input data are processed through layers of artificial neurons connected by weights (Fig. 1). ANNs have been used for classification of ECGs.[3-5] In one study an ANN diagnosed myocardial infarction from ECGs better than a conventional interpretation program[6] and in another study an ANN performed at par with an experienced ECG reader.[7] The purpose of this study was twofold. (1) Develop ANNs that yield high sensitivity for the recognition of right/left arm lead reversal at a very high specificity. (2) Compare the performances of the ANNs to those of 2 widely used interpretation programs.

## Methods

### Study Population

The study was based on 11,432 ECGs recorded on patients who presented at the emergency ward at the University hospital in Lund during 1992 and 1993. The 12-lead ECGs were recorded using computerized electrocardiographs (Siemens-Elema AB, Solna, Sweden).

All recordings were made digitally and averaged heart cycles were calculated. P, QRS, and ST-T measurements used in the criteria and as inputs to the ANNs were obtained from the measurement program of the computerized ECG recorders.

As stated above, the purpose of the study was to develop ANNs that detect right/left arm lead reversal with a very high specificity. Since ANNs learn from example, it was important that no ECG with right/left arm lead reversal was used as an example of correctly recorded ECG and vice versa. Therefore, great care was taken to exclude ECGs with lead reversal as well as those ECGs that were technically deficient. In addition, pacemaker ECGs were excluded since QRS morphology or the ST-T segment are not assessed by the programs in these ECGs. The exclusion process included both visual inspection of the ECGs and computerized methods. Two experienced ECG readers examined the ECGs independently. The computerized methods included histogramming all the variables and a more elaborate method of using an ANN density map.

Technically deficient recordings and pacemaker ECGs were excluded. Based on the visual and computerized analysis, ECGs with suspected electrode misplacement or lead reversal were scrutinized by the 2 experts together. Electrode misplacement or lead reversal was verified in most cases after a comparison between the suspected ECG and a previous or later recording from the same patient found in the clinical database. A total of 423 ECGs were excluded (Table I), which left 11,009 ECGs for the study.

Each of these 11,009 ECGs was used to generate an ECG with right/left arm lead reversal. This was performed computationally by means of inverting lead I, interchanging lead II and III, and interchanging lead aVL and aVR (note that aVF is not affected by the right/left arm lead reversal) (Fig. 2). This yielded exactly the same ECG that would have resulted if the right and left arm electrodes had been switched in the recording situation. Thus, the final material consisted of 22,018 ECGs, divided into 2 groups; 11,009 ECGs recorded with correct electrode placement and 11,009 with right/left arm lead reversal.

The interpretation programs use different criteria depending on the presence/absence of P-waves. Similarly, different ANNs were used if P waves were present or absent. Therefore, only 9,296 pairs of ECGs with P waves could be used to test criteria and ANNs that use P wave information. However, all ECGs were used in the analysis of criteria and ANNs not

requiring P wave data as inputs. This approach is justified since the QRS patterns of right/left arm lead reversal is not dependent on the P waves.

**Conventional criteria**

The performances of the neural networks were compared to those of 2 conventional interpretation programs, namely the GRI program developed at the Glasgow Royal Infirmary[8] and the Marquette program[9]. Both programs use rule based criteria. For example, if P waves are detected the following rule is used in the Marquette program:

**If** QRS axis is between 90 and 270 degrees

**and** P axis is between 90 and 210 degrees,

**then** say "suspect arm lead reversal".

All ECGs were processed by the Glasgow measurement program and these measurements were used as inputs to both interpretation programs. The inputs to the 2 programs are presented in Table II.

**Neural network**

A multilayer perceptron ANN architecture[10] and Langevin updating procedure[11] was used. A more general description of ANN could be found elsewhere.[12] The ANNs consisted of 1 input layer, 1 hidden layer and 1 output layer (Fig. 1). The output unit encodes whether the ECG is correctly recorded (output value = 1) or if a right/left arm lead reversal is detected (output value = 0). The hidden layer contained 4-6 neurons. The number of neurons in the input layer equals the number of input variables. These are presented in Table II for the 6 different networks. P and QRS axis were presented to the networks as sin(axis • _ /180) and cos(axis • _/180). In cases where P or QRS axis were undetermined, both the sin and cos values were set to 0.

Similarly to the criteria, different ANNs were used if P waves were present or not. Two networks used the same inputs as the GRI program while 2 other used the same inputs as the Marquette program. Furthermore, selections of the available inputs were used to find ANNs that yielded high performance.

The data set was divided into 2 parts: a training set and a test set. The training set was used to adjust the connection weights, whereas the test set was used to assess the performance. In order to get as reliable performance as possible a K-fold cross validation was used where the data set was randomly divided into K equal parts. For each of the K different test sets, training was performed on the remaining (K-1)/K parts of the data. We used 3-fold cross validation to decide when to terminate learning in order to avoid "overtraining" and 7-fold cross validation to train the networks and assess their performances. The error estimates in Table III results from 25 independent runs for each type of network.

During the training process the connection weights between the neurons were adjusted using the backpropagation algorithm. The learning rate ($\eta$) had a start value of 0.5. During the training $\eta$ was decreased geometrically every epoch using the following equation:

$$\eta = k\,\eta \quad \text{with } k = 0.998$$

The momentum $\alpha$ was set to 0.7. Updating occurred after each 20 patterns. The Langevin noise was chosen to decrease geometrically from 0.005 and with $k = 0.993$ during the training process. The network weights were initiated with random numbers between -0.025 and 0.025. Parameters which were separate for the different ANNs are presented in Table IV. All calculations were done using the JETNET 3.0 package.[13]

In order to achieve as high specificity as possible with the cost of a lower sensitivity the following 2 methods were used during the training procedure. (1) The correctly recorded ECGs were presented to the ANN typically 50 times more often than the incorrect ECGs. (2) An asymmetric error function was used, i.e. the network was penalised more for a false positive case than for a false negative case.

## Results

The results for the conventional criteria and the ANNs are presented in Table III. All criteria and all ANNs show very high specificity. One of the few correct ECGs classified as right/left arm lead reversal by the ANNs and the Marquette criteria is shown in Figure 3. This ECG was also classified as possible lead reversal by both experts independently. After a comparison with previous ECGs on the patient in a clinical database it was shown that the ECG was correct and the atypical complex in lead I was due to severe lung and heart disease.

In the presence of P waves the conventional criteria show high sensitivity. However, using the ANNs the sensitivity was even higher. Figure 4a shows an example of right/left arm lead reversal which was missed by the GRI criteria. The networks correctly detected the ECG as a lead reversal, which for the experienced ECG reader is obvious. The conventional criteria had a much lower sensitivity in the absence of P wave data; 30.9 and 39.3% for the Marquette and GRI criteria, respectively. The corresponding sensitivity for the ANN which used the same variables as the GRI criteria was 94.5%. One example of lead reversal missed by both criteria and detected by the ANNs is shown in Figure 4b.

In order to check to what extent our database was sufficient in size, the ANNs were trained and tested using half of the material. No significant difference was observed.

## Discussion

One objective of the present study was to develop ANNs with high sensitivity for the recognition of right/left arm lead reversal at very high specificity. The importance of high specificity could be illustrated by the following example. The specificity of the best networks were as high as 99.95%, i.e. 1 false positive ECG of 2000 ECGs. A right/left arm lead reversal occurs only in 1 ECG out of 100-400 depending on the experience of the recording technician. Considering a set of 2000 ECGs, 5-20 of these will have a right/left arm lead reversal and 1 correct ECG will falsely be reported as incorrect using this neural network. With a sensitivity of approximately 95% most of the incorrect ECGs will also be reported. Consequently, most of the ECGs which the neural network classifies as right/left arm lead reversal will actually be a case with lead reversal, i.e. the positive predictive value will be high.

Ideally the network would never report lead reversal for a correct ECG. The network learns from example and it is therefore very important that the training set does not include such examples, i.e. a correct ECG labelled as a lead reversal. This would happen if an ECG with right/left arm lead reversal were included in the data set of 11.009 ECGs labelled as correct. Therefore, great care must be taken in purification of the database. In this study ANNs were used as a complement to the visual examination of the ECG experts in this process.

The performance of an ANN is dependent on the size and composition of the database used for training. Therefore the ECGs of the training set should be representative for routine ECGs processed by computerized ECG-recorders. In order to accomplish this, a database of over 11,000 ECGs was selected. Training and testing ANNs using only half the material did not impair the performance, which shows that the size of the material was sufficient. The ECGs used to train the ANNs were recorded at an emergency ward, i.e. in a clinical setting were the ANNs could be of greatest help.

A second purpose of the study was to compare the performances of the ANNs to those of 2 widely used interpretation programs. Both programs and ANNs showed very high specificities but sensitivities were much higher for the ANNs. Considering that over 100 million ECGs are recorded annually in the US[14] and probably another 200 million in the rest of the world, this difference in sensitivity could result in 100,000-400,000 right/left arm lead reversals being detected by ANNs but not by conventional interpretation programs. It has been shown that even cardiologist fail to recognize right arm/leg lead reversal.[2] The right/left arm lead reversal is probably also overlooked by many ECG readers. Therefore, the detection of lead reversal in interpretation programs is a very important type of quality control.

## References


1 Guijarro-Morales A, Gil-Extremera B, Maldonado-Martin A. ECG diagnostic errors due to improper connection of the right arm and leg cable. *Int J Cardiol* 1991;30:233-235.

2 Haisty WK Jr, Pahlm O, Edenbrandt L, Newman K. Recognition of electrocardiographic electrode misplacements involving the ground (right leg) electrode. *Am J Cardiol* 1993;71:1490-1495.

3 Edenbrandt L, Devine B, Macfarlane PW. Neural networks for classification of electrocardiographic ST-T segments. J *Electrocardiol* 1992;25:167-173.

4 Edenbrandt L, Devine B, Macfarlane PW. Classification of electrocardiographic ST-T segments - human expert versus artificial neural network. *Eur Heart J* 1993;14:464-468.

5 Bortolan G, Degani R, Willems JL. Neural networks for ECG classification. In: *Computers in Cardiology 1990.* Los Alamitos, CA: IEEE Computer Society Press; 1990:269-272.

6 Hedén B, Edenbrandt L, Haisty WK Jr, Pahlm O. Artificial neural networks for the electrocardiographic diagnosis of healed myocardial infarction. *Am J Cardiol* 1994;74:5-8.

7 Reddy MRS, Edenbrandt L, Svensson J, Haisty WK, Pahlm O. Neural network versus electrocardiographer and conventional computer criteria in diagnosing anterior infarct from the ECG. In: *Computers in Cardiology 1992.* Los Alamitos, CA: IEEE Computer Society Press; 1992:667-670.

8 Macfarlane PW, Lawrie TDV. *Comprehensive electrocardiology Volume 3.* Oxford: Pergamon Press Inc; 1989:1529-1530.

9 Physicians Guide to Marquette Electronics Resting ECG Analysis. Marquette Electronics Milwaukee. WI. 1988: 53.

10 Rumelhart DE, McClelland JL, eds. *Parallell distributed processing. Volumes 1 & 2.* Cambridge, MA: MIT Press; 1986.

11 Rögnvaldsson T. On Langevin updating in multilayer perceptrons. *Neural computation* 1994;6:916-926.

12 Hertz J, Krogh A, Palmer RG. *Introduction to the theory of neural computation.* Redwood City, CA: Addison-Wesley; 1991.

13 Peterson C, Rögnvaldsson T, Lönnblad L. JETNET 3.0 - A versatile artificial neural network package. *Comp Phys Comm* 1994;81:185-220.

14 Drazen E, Mann N, Borun R, Laks M, Berson A. Survey of computer-assisted electrocardiography in the United States. *J Electrocardiol* 1988;21:S98-S104.

## Figure Legends

**Figure 1** Schematic diagram of a neural network with 15, 5 and 1 neurons in the input, hidden and output layer, respectively. The neurons are connected to each other as illustrated by the lines.

**Figure 2** A correctly recorded ECG (left) with the corresponding right/left arm lead reversal (right). The result of the computer generated lead reversal is the same as an electrode misplacement namely an inverted lead I, a switch of lead II and III and a switch of lead aVR and aVL.

**Figure 3** A correctly recorded ECG from a 50-year old patient with severe lung and heart disease. The ECG was falsely classified as right/left arm lead reversal by the ANN and the Marquette program.

**Figure 4** Two ECGs with right/left arm lead reversal, 1 with P waves present (a) and 1 without P waves (b). Both ECGs were detected by the ANNs, while the GRI criteria missed them and the Marquette criteria missed the ECG without P waves.

**Table I:** Reasons for exclusion of ECGs.

| Reason for exclusion | Number of ECGs |
|---|---|
| Right/left arm lead reversal | 47 |
| Right arm/foot lead reversal | 32 |
| Left arm/foot lead reversal | 8 |
| Precordial lead misplacement | 25 |
| Pacemaker ECG | 197 |
| Technically deficient ECG | 114 |
| **Total** | **423** |

**Table II:** Measurements used in the GRI and Marquette criteria and in the ANNs.

| | waves prese | | | ot present | |
|---|---|---|---|---|---|
| **Measurement** | **GRI** | **Marquette** | **Best ANN** | **GRI** | **Marquette** |
| P axis | x | x | x | | |
| P sum (I, $V_6$) | | | x | | |
| P+ , P– ($V_6$) | x | | | | |
| QRS axis | x | x | x | x | x |
| QRS area (I) | x | | x | x | |
| QRS area ($V_6$) | x | | x | | |
| Q, R ampl (I) | x | | | x | x |
| S, R´, ST, T+, T- ampl (I) R, S ampl ($V_6$) | x | | | x | |
| QRS PP ($V_6$) | x | | | | |
| QRS ampl (35, 40, 45), T sum (I, $V_6$) | | | x | | |
| ST, T+, T– ampl ($V_6$) | x | | | x | |

+ = maximal positive amplitude; – = maximal negative amplitude; PP = peak-to-peak amplitude; sum = maximal positive amplitude – |maximal negative amplitude| ; QRS ampl (35, 40, 45) = amplitude 35, 40 and 45 ms after QRS onset;

**Table III:** Results for the GRI and Marquette criteria and for the ANNs. The results for the ANNs are presented as mean ± standard deviation.

| Criteria/ANN | Specificity | Sensitivity |
|---|---|---|
| **P waves present** | | |
| GRI criteria | 100.00 | 84.02 |
| ANN using GRI measurements | 99.94 ± 0.026 | 98.7 ± 0.12 |
| Marquette criteria | 99.87 | 93.91 |
| ANN using Marquette measurements | 99.92 ± 0.022 | 94.8 ± 0.26 |
| Best ANN | 99.95 ± 0.017 | 99.11 ± 0.080 |
| **P waves not present** | | |
| GRI criteria | 100.00 | 39.34 |
| ANN using GRI measurements | 99.92 ± 0.023 | 94.5 ± 0.52 |
| Marquette criteria | 99.91 | 30.93 |
| ANN using Marquette measurements | 99.90 ± 0.028 | 63 ± 2.6 |

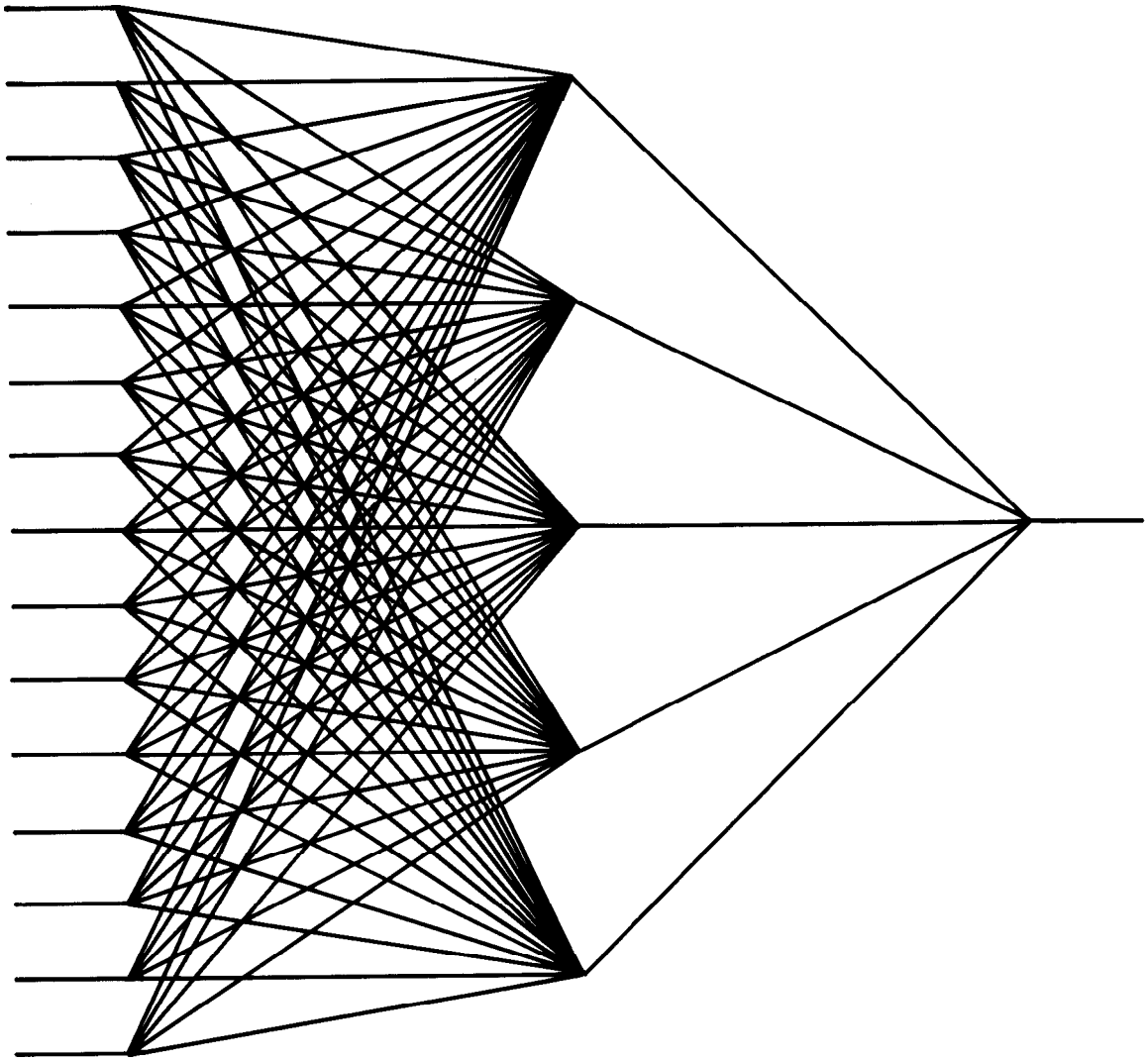
**Table IV:** Network structure for different input data set.

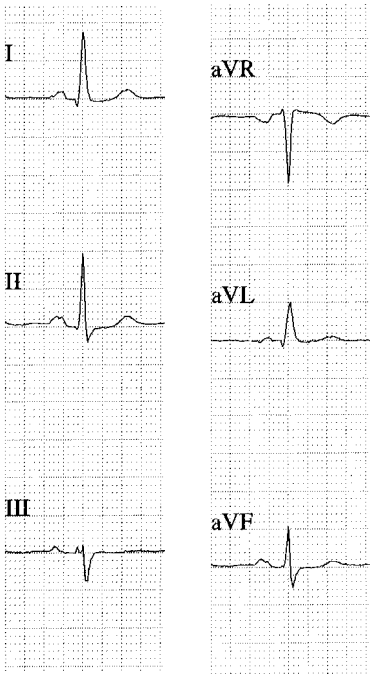
| Neural network | ber of neu | | | SR | AE | $Err_{min}$ | <Epochs> |
|---|---|---|---|---|---|---|---|
| | Input | Hidden | Output | | | | |
| **P waves present** | | | | | | | |
| ANN using GRI measurements | 21 | 5 | 1 | 40:1 | 1 | 0.020 | 148 |
| ANN using Marquette measurements | 4 | 4 | 1 | 70:1 | 1 | 0.030 | 325 |
| Best ANN | 16 | 5 | 1 | 20:1 | 10 | 0.022 | 125 |
| **P waves not present** | | | | | | | |
| ANN using GRI measurements | 15 | 5 | 1 | 70:1 | 1 | 0.030 | 170 |
| ANN using Marquette measurements | 4 | 6 | 1 | 50:1 | 5 | 0.075 | 500 |

SR = sampling ratio between correctly recorded ECGs and ECGs with right/left arm lead reversal; AE = The ß-parameter in the asymmetric error function (AE=1 is the normal squared errorfunction); $Err_{min}$ = Error level in the training set at which the training is stopped; <Epochs> = The average number of epochs, where 1 epoch is completed when the same number of patterns have been presented to the network as the number of patterns in the training set;

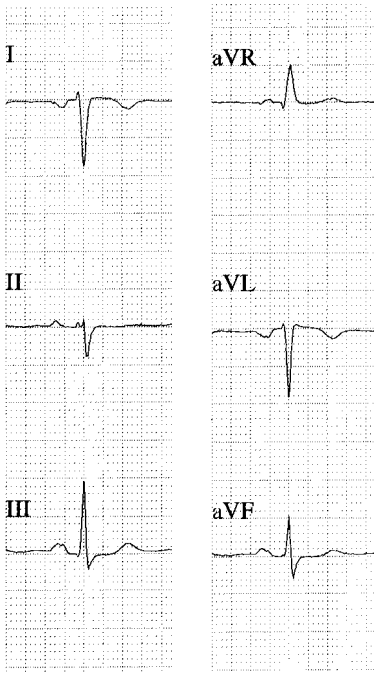
## Appendix

Very often measurements are redundant in the sense that they represent physical quantities that depend upon one another. This is the case for electrocardiographic leads. Hence the different lead measurements could serve as mutual consistency checks. This can conveniently be performed within the ANN picture as follows: Map all variables onto themselves with 1 hidden layer. After training the network computes the deviations between input and output values. Ideally these deviations should all be zero or at least a small number with similar magnitude. If some measurements are wrong or if data points originate from different dynamics those extents will exhibit large deviations. In Figure 5 an example is shown where the QRS area is approximately determined using QRS amplitudes from lead I. In this specific example a linear network was used, meaning that the QRS area is written as a linear combination of all the amplitudes. As can be seen from Figure 6 almost all QRS areas have small errors. The 10-20 ECGs with highest errors were checked by the ECG experts. In this example pacemaker ECGs (excluded), technically deficient ECGs (excluded) and bundle branch blocks (not excluded) were found.

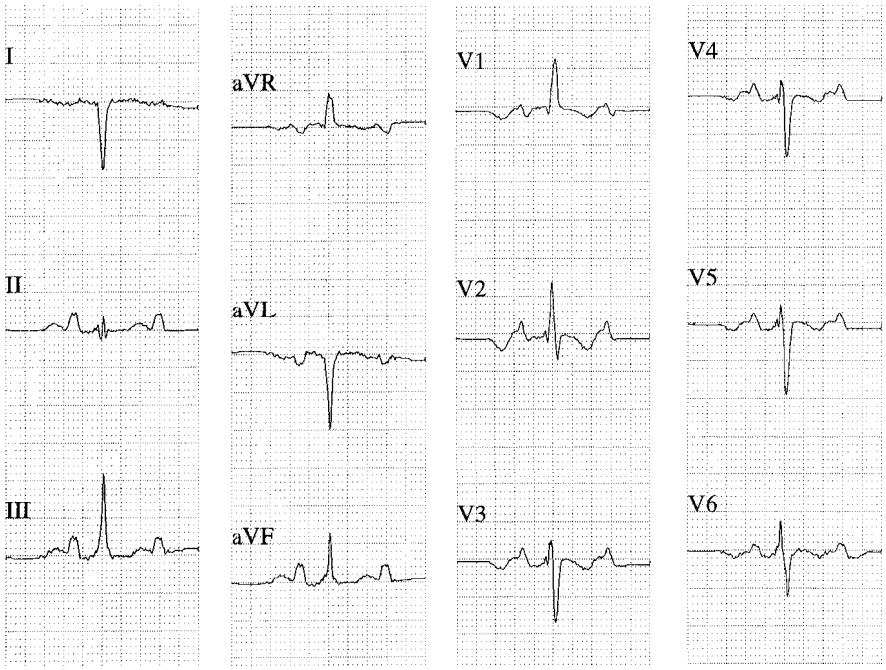

Input layer
Hidden layer
Output layer

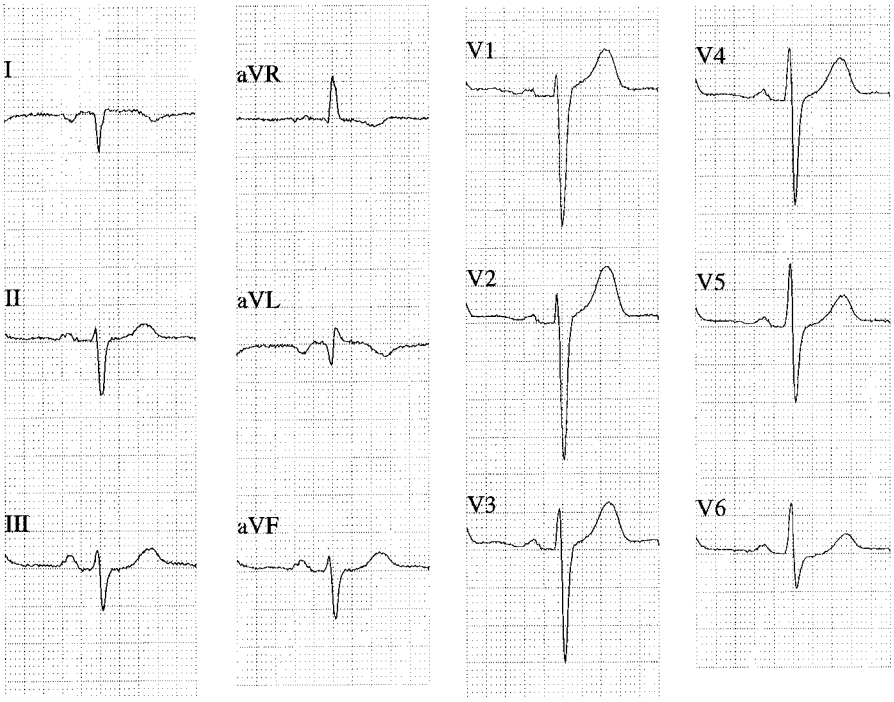


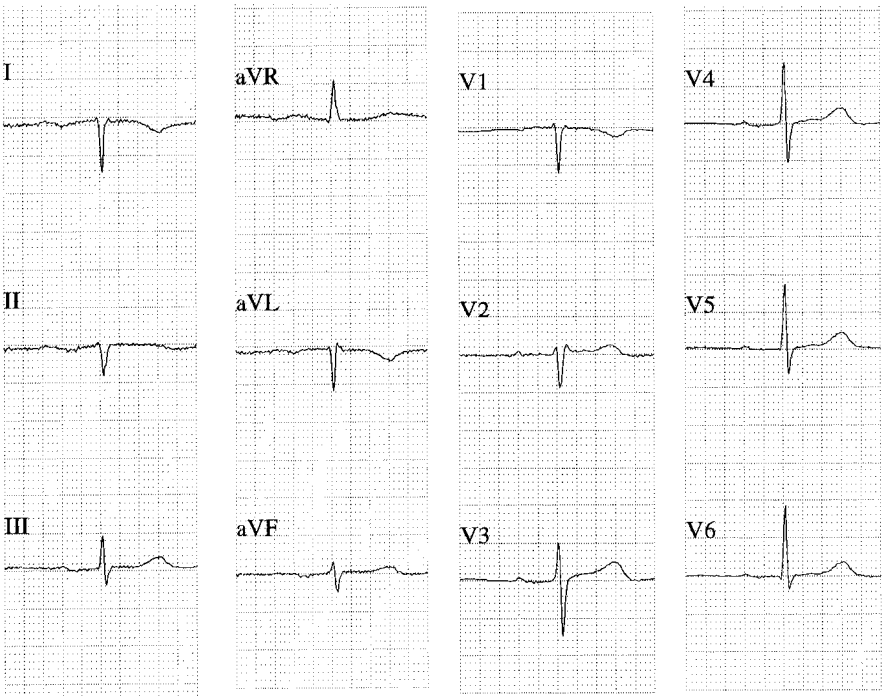

**Correctly recorded ECG**




**Right/left arm lead reversal**


I
II
III
aVR
aVL
aVF
V1
V2
V3
V4
V5
V6


I
II
III
aVR
aVL
aVF
V1
V2
V3
V4
V5
V6


I
II
III
aVR
aVL
aVF
V1
V2
V3
V4
V5
V6