Applications of Artificial Neural Networks for ECG Signal Detection and Classification

Yu Hen Hu, PhD, Willis J. Tompkins, PhD, José L. Urrusti, MS, and Valtino X. Afonso, MS

Abstract: The authors have investigated potential applications of artificial neural networks for electrocardiographic QRS detection and beat classification. For the task of QRS detection, the authors used an adaptive multilayer perceptron structure to model the nonlinear background noise so as to enhance the QRS complex. This provided more reliable detection of QRS complexes even in a noisy environment. For electrocardiographic QRS complex pattern classification, an artificial neural network adaptive multilayer perceptron was used as a pattern classifier to distinguish between normal and abnormal beat patterns, as well as to classify 12 different abnormal beat morphologies. Preliminary results using the MIT/BIH (Massachusetts Institute of Technology/Beth Israel Hospital, Cambridge, MA) arrhythmia database are encouraging. **Key words:** artificial neural networks, adaptive multilayer perceptron structure, QRS detection and classification.

Automating electrocardiographic data monitoring and interpretation is a challenging problem that has made significant progress in recent years. However, the performance levels achieved so far leave room for significant improvement. For example, good contemporary QRS complex detectors correctly recognize more than 99.3% of the beats in a standard database. The misclassifications represent an average of 33 false detections per patient per hour, which is not acceptable in many applications.[1]

Here, we explore the potential applications of an emerging, unconventional computation model called the artificial neural network (ANN) to enhance the performance of electrocardiographic signal detection and classification. An ANN model distinguishes itself from conventional digital computers in that it is a parallel, distributed, nonlinear computing network that mimics the information processing structure of a biologic neural system. Some have argued that the ANN structure is superior to existing digital computer architectures in performing many perceptual-related tasks, such as pattern classification and nonlinear dynamic system modeling. While current ANN structures are simulated using existing digital computers, research and development efforts are underway to build hardware ANN architectures that exhibit massive parallelism and real-time processing capability.

From among the numerous paradigms of the ANN,[2] we focus on a model called the feed forward multilayer perceptron (MLP), which has found numerous applications in pattern classification, time-series modeling, nonlinear control, and other areas. We exploit potential applications of the MLP architecture to perform two important electrocardiographic processing tasks: QRS complex detection and QRS beat pattern classification. For QRS complex detection, we used an adaptive MLP structure to model

From the Department of Electrical and Computer Engineering, University of Wisconsin-Madison, Madison, Wisconsin.

Reprint requests: Willis J. Tompkins, PhD, Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 1415 Johnson Drive, Madison, WI 53706-1691.

the nonlinear background noise which, for the purpose of QRS detection, includes P and T waves, so that the presence of QRS complexes is emphasized. For QRS complex beat classification, we trained an MLP classifier with QRS beat patterns and classified each new pattern presented to the ANN as normal or into 1 of 12 possible abnormal beat categories. We compared the results with those for a nearest neighbor classifier and found that the MLP classifier performed significantly better.

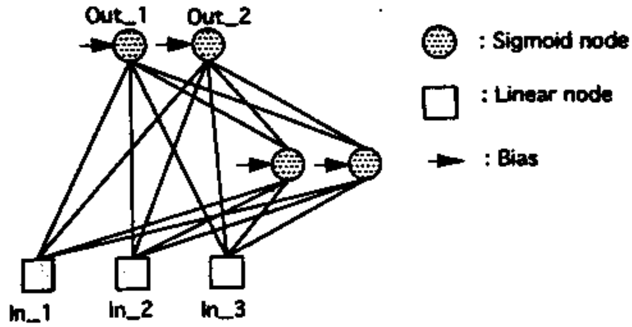

Fig. 2. A 3-2-2 configured multilayer perceptron with direct connection between input and output.

Multilayer Perceptron Artificial Neural Network

Recently, the ANN has received unprecedented attention in many research disciplines.[3–6] The ANN's parallel, distributive computational structure is reminiscent of the human neural system. In an ANN structure, many simple, nonlinear processing elements, called neurons, are interconnected via weighted synapses to form a network. Figure 1 shows that the function of each neuron is to compute a weighted sum of all synapse inputs, subtract the sum from a predefined bias, and pass the result through a nonlinear sigmoidal (threshold) function whose output ranges between 0 and 1. More specifically, the ith output activation a_i is a sigmoid function (defined in equation 1 of a so-called net function u_i, which, in turn, is a weighted sum of all the inputs x_j and a bias θ_i [compare equation 2]).

$$(1) \qquad a_i = f(u_i) = \frac{1}{1 + \exp(-u_i)}$$

$$(2) \qquad u_i = \sum_{j=1}^{N} w_{ij}x_j + \theta_i$$

If neurons are grouped in layers with weighted synapses interconnecting only neurons in successive layers, the ANN structure is called a multilayer perceptron model (Fig. 2). An MLP model is the most popular and most extensively studied ANN model. An MLP consists of an input layer and an output layer, with one or more hidden layers in between. Usually, the input units only hold the input signal without any processing. Computation occurs in neurons in the hidden layers and the output layer. Thus, the structure depicted in Figure 2 is called a two-layer MLP structure.

An MLP can be regarded as a multidimensional nonlinear mapping function from the input space to a hypercube. With a sufficient number of synapse weights, an MLP is capable of approximating any nonlinear functional mapping to arbitrary accuracy. In the QRS complex detection application, we use the MLP to model the background noise process. From a different perspective, since each output is a sigmoidal function, the MLP can also be regarded as a nonlinear discriminating function that distinguishes whether the input vector belongs to a class (eg, has an output value near 1) or not (has an output value near 0). In this way, an MLP structure can be used as a pattern classifier.

The application of an ANN usually consists of a training phase and a testing phase. During the training phase, patterns (samples) are applied at the input and the corresponding desired outputs (the teaching output) are presented to the MLP classifier. A training algorithm is executed to adjust the weights (w_{ij}) and bias (θ_i) so that the actual output of the MLP best matches the desired output. This training process continues until the user is satisfied with the performance. Upon completion of the training phase, a set of test samples, which are not part of the training sample set, are applied to the trained MLP classifier to test whether it is able to generalize reasonably well. The training of an MLP is usually accomplished

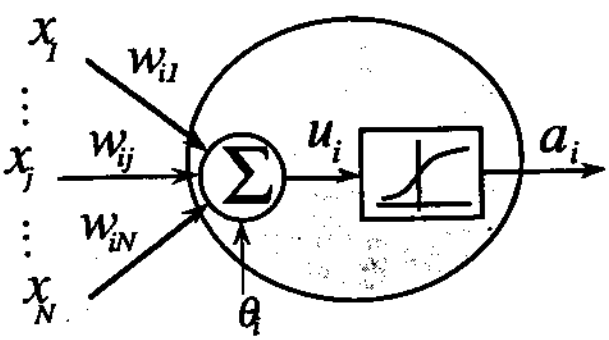

Fig. 1. Functional description of a single neuron.

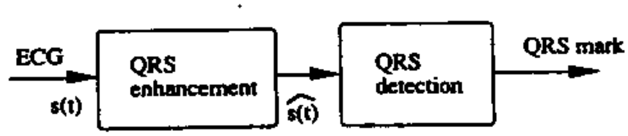

Fig. 3. A QRS detection system.

using a stochastic gradient-based algorithm called the (error) backpropagation algorithm.

Application of the Multilayer Perceptron to QRS Complex Detection

Problem Statement

Reliable detection of the QRS complex enables continuous vital sign monitoring and rhythm analysis of the electrocardiographic signal. Our past efforts in QRS detection algorithm development have yielded satisfactory results for detecting the QRS in noiseless or low-noise environments.[1,7–12] The objective here is to study whether using an ANN-based approach will enable us to further improve the performance of the QRS detector. Specifically, we focused on the ability to adaptively remove nonlinear, time-varying background noise.

Approach

Figure 3 depicts a block diagram of the QRS detection algorithm. The detection is carried out in two phases. In the QRS enhancement phase, the background noise, including P and T waves, is removed so that the signal (QRS complex) to noise ratio is enhanced. The enhanced QRS complex signal then passes through the QRS detection block to detect and mark the position of the QRS complex.

For the purpose of QRS enhancement, we assume the electrocardiographic signal $x(t)$ as the superposition of the QRS complex $s(t)$ and background noise $n(t)$. That is $x(t) = s(t) + n(t)$. The objective of QRS enhancement is to find an output $\hat{s}(t)$ that is a (deformed) approximation to $s(t)$. This is found by subtracting a model of the background noise process $\hat{n}(t)$ from $x(t)$ as depicted in Figure 4. Using an MLP model, $\hat{n}(t)$ is estimated as a nonlinear combination of the p past samples of $x(t)$, $\{x(t-i);\ 1 \le i \le p\}$. Using the backpropagation training algorithm, we adaptively update the weights of the MLP model so that the mean-square prediction error $E\{|x(t) - \hat{n}(t)|^2\}$ is minimized. Using the fact that a QRS complex has a high slope that cannot be predicted well with only p past samples, the prediction error $[x(t) - \hat{n}(t)]$ will exhibit a significantly large error when a QRS complex is present and will be relatively small in its absence. Thus, the error signal can be regarded as a deformed approximation of the QRS complex signal $s(t)$. This signal is then sent to the QRS detection block for further processing. The details of the QRS detection block are omitted here due to space limitations. They can be found in Hamilton and Tompkins.[1]

Experimental Results

We conducted a preliminary experiment using data records 105 and 108 of the MIT/BIH arrhythmia database (Massachusetts Institute of Technology, Cambridge, MA). We chose these records because they are particularly noisy and cannot be handled well with existing methods.

We used a two-layer MLP ANN with a 6-3-1 configuration. We selected the input window size p (6) and the number of hidden units (3) empirically. We trained the MLP using the backpropagation algorithm during an initial training period. The weights

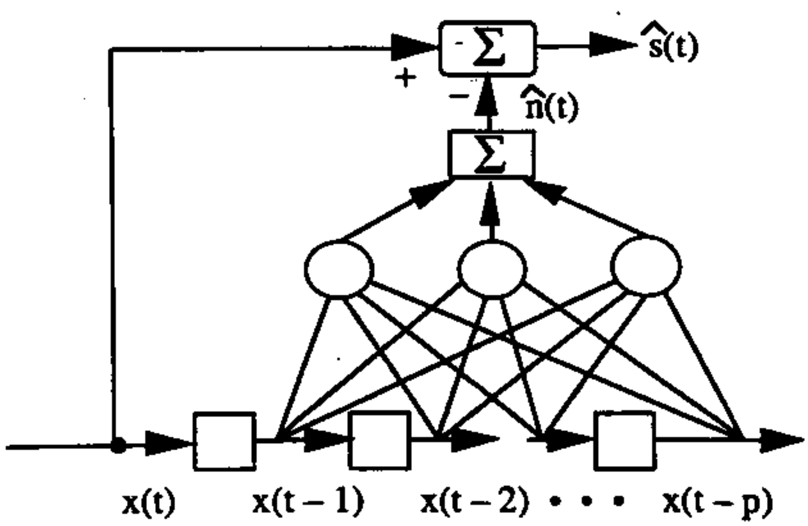

Fig. 4. Multilayer perceptron-based QRS enhancement.

Table 1. Results of the Evaluation of Three QRS Detection Algorithms Using Record 105 of the MIT/BIH Database With a Total of 2,572 Heartbeats

Filtering Methods	False Positive (Beats)	False Negative (Beats)	Total Failed Detection (Beats)	Failed Detection Rate (%)
Artificial neural network adaptive, nonlinear	10	4	14	0.50
Linear adaptive filtering	40	22	62	2.40
Bandpass filtering	67	22	89	3.46

were adaptively updated during testing for each new incoming data point.

In Tables 1 and 2, we compare the QRS complex detection rates of two existing QRS enhancement methods, the linear adaptive filtering method and the bandpass filtering method, to the adaptive MLP method using complete data records 105 and 108 of the MIT/BIH database, which include 2,572 and 1,763 beats, respectively. For data record 105, the ANN-based filtering method had only 10 false positive (FP) and 4 false negative (FN) beats out of 2,572 beats in the record. This corresponds to a total failure rate of 0.5% ([10 + 4)/2,572] × 100), which is much better than the failure rate of the bandpass filtering method (3.5%) and the linear adaptive filtering method (2.4%). For record 108, the ANN method had 25 FP and 16 FN beats, which is a 2.3% failure rate. The failure rates of the bandpass filter and the linear adaptive filter approaches were also higher, 12.5% and 4.4%, respectively. The potential advantages of an MLP-based approach are quite clear from these preliminary experiments.

Applications of the Artificial Neural Network to Electrocardiographic Beat Classification

QRS beat classification is a crucial task in electrocardiographic diagnosis. Many current beat pattern classification methods rely on various features extracted (measured) from the electrocardiographic waveform. For example, such measurements as the QRS width and amplitude are often used as part of a measurement matrix of extracted features. The actual QRS pattern (the raw data) is currently used in template matching sorting techniques.

To date, several researchers have reported attempts at using an ANN to classify electrocardiographic beats. Yeap et al.[13] proposed to use the amplitude of the QRS, the offset of the QRS, the T-wave slope, and the prematurity as the inputs to an ANN structure with 20 hidden units. On the American Heart Association database, they reported a 67.6% classification rate on the test set. Linnenbank et al.[14] used a three-layer ANN model to classify 62-lead ventricular tachycardia QRS integral map patterns. Pattern classification was obtained with reference to a database of 18 different paced QRS integral map patterns. Each reference pattern corresponds to activation onset in a localized left ventricular segment. In comparison to visual classification by an experienced human observer, 92% of the tested 77 ventricular tachycardia QRS integral map patterns were correctly classified. Tsai et al.[15] used power spectral density of the electrocardiographic signal as the input feature to classify five different types of normal and abnormal electrocardiographic beats. Lee[16] reported the use of a higher-order ANN model for electrocardiographic classification. We previously proposed a method to train the network with shifted versions of the electrocardiographic signal.[17] The test results show that the network trained by this method is more robust for time-shifted versions of the same type of electrocardiographic signals. Bortolan et al.[18] reported on the use of input features, including both morphology signals and multiplicative terms, with an order as high as 39 and a total of 780 input features.

Table 2. Results of the Evaluation of Three QRS Detection Algorithms Using Record 108 of the MIT/BIH Database With a Total of 1,763 Heartbeats

Filtering Methods	False Positive (Beats)	False Negative (Beats)	Total Failed Detection (Beats)	Failed Detection Rate (%)
Artificial neural network adaptive, nonlinear	25	16	41	2.32
Linear adaptive filtering	58	20	78	4.42
Bandpass filtering	199	22	221	12.54

In the next section, we apply an MLP classifier to directly classify the QRS complex wavelet patterns into the normal class or into 1 of 12 abnormal classes. We compare the results with a nearest neighbor classifier, a widely used classification method.

Methods

We used an MLP structure as a pattern classifier. The input to the MLP is the original QRS beat patterns with appropriate linear amplitude scaling. Each output is associated with a particular class. If the output corresponding to a class is maximum among all other outputs for a given QRS beat pattern, that beat is classified into that class. We used different numbers of hidden units and outputs for different experiments. We added direct connections from the input nodes to the output nodes in this application as we found that it helped in improving the classification performance.

We extracted 6,474 QRS complex templates classified into 13 beat types from the MIT/BIH arrhythmia database. Each template contains 51 samples centered on the annotation fiducial mark in the record. Since the sample rate of the database is 360 Hz, each QRS template represents a time segment of about 142 ms. Figure 5 illustrates several of these data records. Data in each record are taken from the same patient. From Figure 5 we can make several observations. First, QRS beats within each record are relatively clustered. Second, the morphology of waveforms in the same class (ie, with the same beat annotation), but from different records (patients), can exhibit significant variations. For example, left bundle branch block morphologies from different patients are quite different. Third, QRS beats from different classes and different patients may look very similar. For example, nodal premature beat QRS complexes are very similar in morphology to normal QRS beats. In this study, we did not use beat-to-beat time intervals to help in the classification. We used only the beat morphologies. Figure 5 shows an overlay of the 50 nodal premature beats selected from the MIT/BIH database for this study. We did not normalize the amplitudes, remove baseline offsets, or adjust time alignment of any of the beat sets that we used. They were simply removed from the database using the annotation location as a fiducial point.

Three-way Cross Validation. To estimate the classification rate accurately, we adopted the cross-validation method that is designed to minimize the variations due to imperfect random sampling of finite-size data samples. We partitioned each class of the data set randomly into three disjoint subsets of approximately equal size denoted by, say, A, B, C. The training and testing was performed three times each with each one of the three subsets as the test set and the other two as the training set. We did the actual partition using MATLAB (The Mathworks, Inc., Natick, MA) software. Table 3 summarizes the results.

Preprocessing. For ANN applications it is customary to scale the input data so that it has a dynamic range between −1 and +1. Since most data entries lie between 800 and 1,600, we scaled each entry by 0.001 and then subtracted 0.8 from the result, thereby, placing each scaled data entry in the range of 0 to 0.8.

Nearest Neighbor Classifier. We used a simple nearest neighbor classifier as the baseline comparison. First, we computed the mean (centroid) of the samples in each class. Then we calculated the Eu-

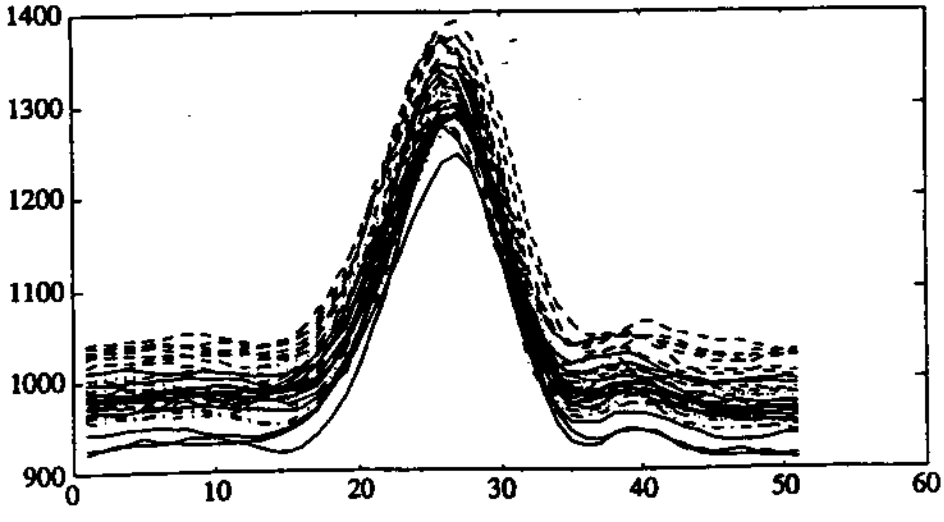

Fig. 5. Overlay of 50 nodal (junctional) premature beats used in this study.

Table 3. Distribution of Samples in Each Class and Each Partition

Beat Type	Number of QRS Templates	No. of Records	Set_a	Set_b	Set_c
Normal	3,891	39	1,299	1,299	1,293
Left bundle branch block	200	4	67	67	66
Right bundle branch block	300	6	101	101	98
APC	383	8	128	128	127
ABERR	50	1	17	17	16
Nodal premature beat	50	1	17	17	16
PVC	1,000	20	334	334	332
Fusion	100	2	34	34	32
Flwave	50	1	17	17	16
NESC	50	1	17	17	16
VESC	50	1	17	17	16
Pace	200	4	67	67	66
PFUS	150	3	51	51	48
Total	6,474	—	2,166	2,166	2,142

clidean distances between a test vector x and each of these centroids $\{w_i;\ 1 \le i \le \#\text{—classes}\}$. The test vector was classified into class i^* if

$$(3) \quad i^* = \arg_i^{\min} d(x, w_i) = \arg_i^{\min} \| x - w_i \|.$$

We used no elaborate distance measurement or fine tuning of the centroid locations. Three-way cross validation was also adopted in this technique to provide a fair basis for comparison.

Training and Testing. We used a public-domain backpropagation training package called ASPIRIN v.6 (MITRE Co., Boston, MA) as the training tool. Based on a given network configuration file, ASPIRIN generates optimized C language code on the target platform, which, in this experiment, was a Hewlett-Packard 720 precision architecture RISC workstation. We adjusted the learning rate α and the moment μ for each training set empirically to maximize the classification rate. We used a maximum of 2,000 epochs to limit the number of training iterations. Each epoch consisted of presenting all of the training samples in the training set once. The training was stopped when either all training samples were correctly classified or when the iteration limit was reached.

Upon completion of the training phase, we extracted the weights of the trained network and fed them into a MATLAB program to compute the classification rate on the test set on a class-by-class basis. Even in the first experiment where only a normal or abnormal decision must be made, it is important to know how each individual abnormal class performs with the given classifier.

Experimental Results

We performed two experiments. In the first experiment, we attempted to classify all 13 classes simultaneously using an MLP structure configured as 51-40-13. Table 4 summarizes the results. The classification rate in this experiment was about 65%. Although this is low, it is still much better than the nearest neighbor classifier results that achieved, on average, less than 30% classification rate.

In the second experiment, we first classified each beat pattern as a normal or a abnormal beat (two outputs). We accomplished this with a 51-25-2 MLP structure. The classification rate, which is tabulated in Table 5, is typically about 90%. In a separate MLP network with a configuration of 51-30-12, we used 12 outputs, each denoting 1 of the 12 abnormal classes. A normal beat applied to this network would have a target value of all zeroes. This second MLP network categorized the beats identified as abnormal by the preprocessor MLP into 1 of the 12 abnormal classes. The results are tabulated in Table 6. Upon completion of the training and testing of both MLP structures, we cascaded the two networks such that the output of the first MLP, which indicates when a beat is abnormal, was applied to mask out the output of the second network. In other words, if the first network determined that a given QRS complex was normal, then all of the outputs of the second network were reset to zero. Only when the QRS beat was deemed to be an abnormal beat was the output of the second network allowed to pass through the mask and indicate an abnormal class. Using this composite MLP classifier, the total classification rate improved from 65.5% in the first experiment to 84.5%. Table 7 shows the results.

Conclusion

We trained an ANN MLP classifier with QRS beat patterns and classified each new pattern presented to the ANN as normal or into 1 of 12 possible abnor-

Table 4. Classification of All 13 Classes Using a 51-40-13 Multilayer Perceptron

Beat Type	Trial_a	Trial_b	Trial_c	Average
Normal	60.89	64.36	47.80	57.70
Left bundle branch block	88.06	98.51	93.94	93.50
Right bundle branch block	84.16	90.10	93.88	89.33
APC	44.53	55.47	60.63	53.52
ABERR	82.35	88.24	87.50	86.00
Nodal premature beat	88.24	100.00	100.00	96.00
PVC	72.75	74.85	79.22	75.60
Fusion	79.41	82.35	78.13	80.00
Flwave	82.35	94.12	75.00	84.00
NESC	88.24	100.00	100.00	96.00
VESC	100.00	82.35	100.00	94.00
Pace	95.52	92.54	89.39	92.50
PFUS	45.10	62.75	66.67	58.00
Total	65.74	69.94	60.78	65.51
Nearest neighbor	24.79	26.08	25.92	25.60

Table 5. Classification of Normal versus Abnormal Beats

Classes	Trial_a	Trial_b	Trial_c	Average
Normal	93.69	90.84	98.69	94.40
Abnormal	87.77	88.34	78.23	84.83
Total	91.32	89.84	90.58	90.58
Nearest neighbor	71.14	73.91	73.90	72.98

Table 6. Classification Among All Abnormal Beats

Beat Type	Trial_a	Trial_b	Trial_c	Average
Left bundle branch block	100.00	97.01	86.36	94.50
Right bundle branch block	96.04	95.05	95.92	95.67
APC	59.38	64.84	64.57	62.92
ABERR	94.12	88.24	87.50	90.00
Nodal premature beat	100.00	100.00	100.00	100.00
PVC	76.05	86.23	72.89	78.40
Fusion	94.12	70.59	87.50	84.00
Flwave	76.47	94.12	100.00	90.00
NESC	88.24	100.00	100.00	96.00
VESC	100.00	82.35	100.00	94.00
Pace	91.04	92.54	77.27	87.00
PFUS	60.78	70.59	81.25	70.67
Total	80.28	84.54	79.03	81.29

Table 7. Composite Classification Rate

Beat Type	Trial_a	Trial_b	Trial_c	Average
Normal	93.69	90.84	97.99	94.17
Left bundle branch block	100.00	97.01	86.36	94.50
Right bundle branch block	92.08	91.09	77.55	87.00
APC	30.47	35.16	25.98	30.55
ABERR	94.12	88.24	87.50	90.00
Nodal premature beat	0.00	0.00	0.00	0.00
PVC	75.45	85.63	71.39	77.50
Fusion	82.35	58.82	25.00	56.00
Flwave	70.59	88.24	68.75	76.00
NESC	52.94	82.35	12.50	50.00
VESC	100.00	82.35	100.00	94.00
Pace	91.04	92.54	69.70	84.50
PFUS	47.06	64.71	64.58	58.67
Total	84.72	85.00	83.94	84.55

mal beat categories. For this study, we extracted 6,474 beats having 13 different annotations from the MIT-BIH electrocardiographic database. We presented only the beat morphologies to the ANN. We did not use any features extracted from the signal, such as beat-to-beat time intervals. We classified about 84.5% of the beats correctly based on morphology alone. Although this is a much lower performance than many other published QRS complex classifiers, we expect that we can improve on this result by (1) providing the network with parameters about each beat, such as its temporal relationship with the beats around it and (2) training the network with a larger number of QRS complexes.

References

1. Hamilton PS, Tompkins WJ: Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans Biomed Eng BME-33: 1157, 1986
2. Simpson PK: Artificial neural systems. Pergamon Press, Tarrytown, NY 1990
3. Rumelhart DE, MacClelland JL: Parallel distributed processing: explorations in the microstructure of cognition. Vol. 1–2. MIT Press, 1986
4. Hopfield JJ: Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci USA 79:2554, 1982
5. Kohonen T: Self-organization and associative memory. Springer-Verlag, Berlin, 1984
6. Lippmann RP: An introduction to computing with neural nets. IEEE ASSP Magazine 4:22, 1987
7. Thakor NV, Webster JG, Tompkins WJ: Reliable R-wave detection from ambulatory subjects. Biomed Sci Instrum 14:67, 1978
8. Thakor NV, Webster JG, Tompkins WJ: Optimal QRS detector. Med Biol Eng Comput 21:343, 1983
9. Ahlstrom ML, Tompkins WJ: Automated high-speed analysis of Holter tapes with microcomputers. IEEE Trans Biomed Eng BME-30:651, 1983
10. Pan J, Tompkins WJ: A real-time QRS detection algorithm. IEEE Trans Biomed Eng BME-32:230, 1985
11. Luo S, Tompkins WJ: 12-lead electrocardiogram analysis program. p. 623. In Proceedings of the International Symposium on Measurement Technology and Intelligent Instruments, 1989
12. Hamilton PS, Tompkins WJ: Adaptive matched filtering for QRS detection. p. 147. In Harris G, Walker C (eds): Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 10, 1988
13. Yeap TH, Johnson F, Rachniowski M: ECG beat classification by a neural network. p. 1457. In Pedersen PC, Onaral B (eds): Proceedings of the Annual International Conference of the IEEE Engineering of Medicine and Biology, Vol. 12, 1990
14. Linnenbank AC, Groenewegen AS, Grimbergen CA: Artificial neural networks applied in multiple lead electrocardiography: rapid quantitative classification of ventricular tachycardia QRS integral patterns. p. 1461. In Pedersen PC, Onaral B (eds): Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 12, 1990
15. Tsai YS, Hung BN, Tung SF: An experiment on ECG classification using back-propagation neural network. p. 1463. In Pedersen PC, Onaral B (eds): Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 12, 1990
16. Lee SC: Using a translation-invariant neural network to diagnose heart arrhythmia. p. 2025. In Pedersen PC, Onaral B (eds): Proceedings of the Annual International Conference of the IEEE Engineering Medicine and Biology Society, Vol. 12, 1990
17. Xue QZ, Hu YH, Tompkins WJ: Training of ECG signals in neural network pattern recognition. p. 1465. In Pedersen PC, Onaral B (eds): Proceedings of the Annual International Conference of the IEEE Engineering of Medicine and Biology Society, Vol. 12, 1990
18. Bortolan G, Degani R, Willems JL: Design of neural networks for classification of electrocardiographic signals. p. 1467. In Pedersen PC, Onaral B (eds): Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vol. 12, 1990