SUPERVISED CLASSIFICATION OF ECG USING NEURAL NETWORKS

N. Belgacem[1], M.A Chikh[2], F. Bereksi Reguig[3]

Biomedical Engineering Laboratory, Department of Electronics, Science Engineering Faculty,
Abou Bekr Belkaïd University, B.P.230, Chetouane, Tlemcen, Algeria.
Fax : (213) 43 28 56 85
Email: [1]n_belgacem@hotmail.com,[2] c_wamine@yahoo.fr,[3] f_bereksi@mail.univ-tlemcen.dz

Abstract - In this study, two kinds of neural networks are employed to develop a supervised ECG beat classifier. In order to improve the performance of the MLP classifier for application to ECG signal, the performance is compared to an LVQ neural network classifier. The two classifiers are tested with selected ECG time series and experimental results show that the MLP classifier offers a great potential in the supervised classification of ECG beats.

Keywords- ECG beat classifier, supervised classification, LVQ neural networks.

1. INTRODUCTION

An ECG recording is a measure of the activity of the heart from electrodes placed at specific locations on the torso. A synthesized surface recording of one heartbeat during sinus rhythm can be seen in fig. 1.

Arrhythmia recognition is indeed important for computer assisted automatic diagnosis of cardiac diseases, which can be regarded as a problem of classification of ECG beats. Different clusters in a classifier correspond to different types of ECG beats. Many signal processing techniques have been proposed in the literature for the classification of ECG beats, such as frequency analysis [1-2], template matching [3], and hidden Markov field [4-6].

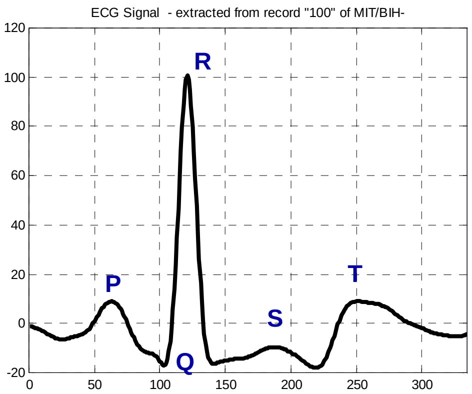

Figure. 1. Synthesized ECG recording for one heartbeat.

The major problems in the automatic classification of ECG beats are: considerable variation of the appearance of normal and abnormal ECG signals, inaccuracies in input feature vectors of the classifier caused by substantial overlapping of the frequencies of the P-wave, T-wave, QRS complex and the noise.

2. DATA MODELING

In this section we describe the principal features extracted from the ECG signal. These features present each heart beat and will be used as the input vector of our neural classifier. We concentrate on the classification of normal and abnormal PVC beats. ECG records of twelve patients were selected from the MIT-BIH arrhythmia database [7] shown in table I. The sampling frequency of the ECG signals in this database is Fs = 360Hz.

TABLE I.

EVALUATION DATA TAKEN FROM THE MIT-BIH ARRHYTHMIA DATABASE.

MIT Records	Normal Beats	Number of PVC
106	1507	520
116	2302	109
119	1543	444
200	1743	826
205	2571	71
208	1586	992
210	2423	194
213	2641	220
219	2082	64
221	2031	396
223	2029	473
228	1688	362
Total	**24146**	**4671**

2.1. Preprocessing

In this study, we concentrate on the classification of the Premature Ventricular Contraction (PVC) beats. The availability of annotated MIT BIH database has enabled the evaluation of performance of the proposed beat classification algorithm. In this database, the analog outputs of the playback unit are filtered to limit analog-to-digital converter saturation and for anti-aliasing, using a bandpass analog filter with a passband from 0.1 to 100 Hz relative to real time. In this study, because of its simplicity and fidelity, an all integer coefficient digital bandpass filter, proposed by Lo et al. [8], was used to remove noise caused by power line interference, respiration, muscle tremors, and spikes. Other types of filters have been developed for this use, such as in [9] and [10]. The integer coefficient bandpass filter was formed by combining a lowpass filter with a highpass filter, both based on a sampling frequency of fs = 360Hz. The transfer function of the lowpass filter is given as:

$$H_L(z)=\frac{1-2z^{-6}+z^{-12}}{1-2z^{-1}+z^{-2}}$$

The 3 dB point is at 20HZ, and the first side-lobe zero amplitude is at 60 Hz. Therefore, power line interference at 60 Hz is completely eliminated, and high frequency muscle tremor noise is minimized, which is predominately a result of the bandlimited (anti-aliased filtered) data in the MIT-BIH arrhythmia database. Once the lowpass filter has removed the high frequency noise. The transfer function of the highpass filter is given as:

$$H_H = z^{-127} - \frac{1}{2^{14}} \qquad \frac{1-2z^{-128}+z^{-256}}{1-2z^{-1}+z^{-2}}$$

Where 2^{-14} is the normalization factor. The cutoff frequency of this filter is at 1 Hz, where the gain is unity. Thus, it successfully removes the drift caused by respiration at about 0.2 Hz.

2.2. QRS detection

As a first step in our analysis we implemented our own QRS detection algorithm [11] to extract a number of features in each pulse including the time and waveform of QRS complex (fig. 2).

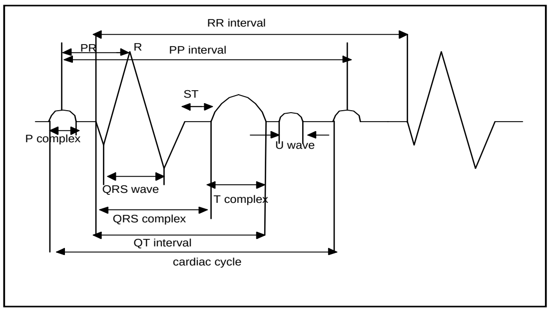

Figure. 2. Detection of different waves in ECG signal.

2.3. Description of data

The QRS complexes were extracted from the band-pass filtered data based on the MIT-BIH arrhythmia database annotations. The QRS segments are obtained as 30 point templates. The position of annotation labels is used to identify the peak of the QRS waveform, and with 15 points on one side and the remaining on the other side with respect of the R peak were picked up to form the template.
The information of each beat is stored as 32-element vector, with the first 30 elements representing the QRS segment, the next two elements representing the temporal parameters such as the instantaneous RR and the width of the QRS complex. The instantaneous RR interval is calculated as the difference between the QRS peak of the present beat and the previous beat.

3. DESCRIPTION AND TRAINING OF THE NEURAL NETWORKS

A set of least object-sensitive ECG features was selected to form the neural network input vector. This strategy aimed to reduce the complexity of the network and facilitate the training phases of the network by employing smaller training data size.

A multi-layer perceptron (MLP) classifier was designed to separate two most common ECG waveforms in the MIT/BIH database. The selected categories were N and PVC beats. The second neural network, with a similar output to the first, was designed to perform a supervised waveforms based on Linear Vector Quantization (LVQ). The number of hidden layer neurons was chosen as ten for the first and 30 for the second, which is small enough for fast training and avoidance of overtraining, yet sufficiently large to give adequate network accuracy [12]. The number of network outputs was selected equal to two in each of the first and the second network. The proposed MLP with a single hidden layer is shown in figure 3.

A target vector was arranged as the desired output for each class. Accompanying each record in the MIT/BIH database in an annotation file in which each heart beat has been identified by expert cardiologist annotators. This annotated information can be employed for designing the target vector and evaluating the classifier performance.

The experimental results of both MLP and LVQ architectures are presented in the following section.

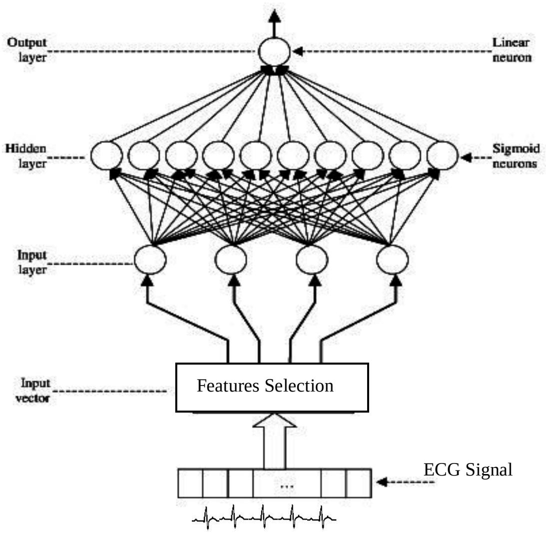

Figure.3. Architecture of the MLP neural network.

As indicated in figure 4, after 09 training epochs, a lower mean square error (MSE) and a smaller gradient were achieved using the training set. Width a smaller size of training set, the network performed better in training but had poor performances when applied to the evaluation or test sets. The two networks were trained with different waveforms extracted from the signal 106, so 1500 and 520 exemplars of normal and PVC beats respectively.

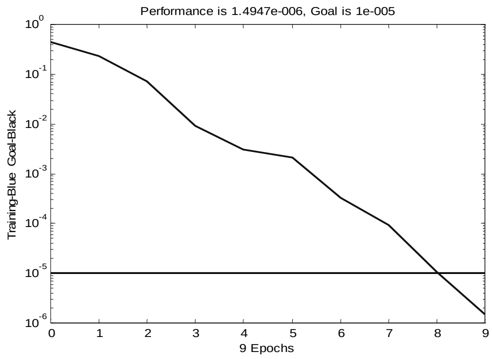

Figure.4 . Training epochs of the MLP Network.

4. RESULTS

The ECG recordings in MIT/BIH database contain a wide range of ECG waveforms; we have chosen a set of data which present only two classes: N and PVC including a total of 10 ECG signal, so 24146 normal and 4671 PVCs.

The results of applying statistical analysis to the classified ECG waveforms by the two neural networks are summarized in Table II and III.

Those classifiers were not designed to detect other abnormalities and showed poor performance in recognizing those beats.

Construction of the input vector based on the extracted features of the signal improves the performance of the classifiers. For each record (as listed in Table II and III), the average recognition rate of the waveforms appearing in the record was found. The overall average recognition rate can be defined as the average of the average recognition rate of the records.

TABLE II.

THE STATISTICAL RESULTS OF THE ECG CLASSIFICATION BY **LVQ NN**.

Record	TP	FP	FN	TN	CC	SE	SP
116	109	616	1	1683	74.39	99.09	73.21
119	444	114	0	1426	94.25	100.00	92.60
200	820	122	37	1619	93.88	95.68	92.99
208	1274	134	94	1450	92.28	93.13	91.54
210	213	499	14	1921	80.62	93.83	79.38
213	418	105	192	2533	90.86	68.52	96.02
219	63	244	9	1835	88.24	87.50	88.26
221	396	53	0	1975	97.81	100.00	97.39
223	493	301	83	1725	85.24	85.59	85.14
228	358	150	6	1536	92.39	98.35	91.10
Total	**4588**	**2338**	**436**	**17703**	**89.00**	**92.17**	**88.76**

TABLE III.

THE STATISTICAL RESULTS OF THE ECG CLASSIFICATION BY **MLP NN**.

Record	TP	FP	FN	TN	CC	SE	SP
116	109	1	0	2299	99.95	100	99.95
119	443	0	1	1541	99.94	99.77	100
200	823	2	295	1316	87.81	73.61	99.85
208	1492	0	83	82	85.25	100	99.05
210	220	14	7	1916	92.65	95.24	97.44
213	213	7	7	2632	99.51	96.81	99.73
219	59	2	5	2078	99.67	92.18	99.9
221	396	831	0	1198	65.73	100	59.04
223	493	301	83	1725	85.24	85.59	85.14
228	359	1	3	1685	99.8	99.17	99.94
Total	**4607**	**1159**	**484**	**16472**	**91.55**	**94.24**	**94.00**

These results show a good performance for normal and abnormal PVC beats classification using neural networks. The average results obtained by the two classifiers were:

89% correct classification, **88.17%** specificity and **92.17**sensitivity for the first classifier,

91.55% correct classification, **94%** specificity and **94.24%** sensitivity for the second classifier.

5. DISCUSSION AND CONCLUSION

The problem of ECG variations among persons with the same cardiological conditions affects the performance of ANN-based cardiac arrhythmia classifiers. In this study, the features were measured from the QRS complex and statistical parameters to cover the important diagnostic information of cardiac events. The features extraction criterion enhanced the reliability and decreased the structural complexity of the ECG classifier. It is also noticed that the MLP classifier results in faster learning, and better classification performance with fewer neurons, as compared with LVQ classifier. However, further investigation is required to a larger class of cardiac arrhythmias and automating the process for microcomputer system implementation.

The results of applying the proposed classifier to the tested records confirmed the objectives of this research, i.e.

- To improve the reliability of the heart beat classifiers,
- To increase the accuracy of diagnosis for classifying arrhythmia cardiac,
- To decrease the structural complexity of the classifier networks.

The design of a perceptron network is constrained completely by the problem to be solved. The number of network inputs and the number of neurons in each layer were constrained by the number of inputs and outputs required by the problem.

REFERENCES

[1] B.H. Hung, Y. S. Tsai and T. H. Chu, "FFT algorithm for PVC detection using IBM PC", *Proc. IEEE Eng. Med. Biol. Soc. 8th annual International conf. pp.292-295, 1986.*

[2] M.A. Chikh, N.Belgacem, F. Breksi-Reguig "A PVCV beat recognition with Fourier analysis and neural network", *Séminaire National sur l'Automatique et les Signaux SNAS'02, Annaba : 27-28 octobre 2002.*

[3] S.A. Caswel, K.S Kluge and CM Chiang, "Pattern recognition of cardiac arrhythmias using two intracardiac channels", C*omputer in Cardiology, pp. 549-554, 1993.*

[4] D. A. Coast, R. M. Stern, and G. G. Cano, "An approach to cardiac arrhythmia analysis using hidden Markov models", *IEEE Tran. Biom. Eng, vol. 37; no. 9, pp. 826-836, September 1990.*

[5] N .V. Thakor and Y. S. Zhu, "Application of cardiac filtering to ECG analysis: Noise calcellation and arrhythmia detection", *IEEE Tran. Biom. Eng, vol. 39; no. 8, pp. 785-794, August 1991.*

[6] W. H. Chang, K.P. Lin and S. Y. Tseng, "ECG analysis based on Hilbert transform in ECG diagnosis", *Proc. IEEE Eng. Med. Biol. Soc, 10th annual International conf. pp.36-37, 1988.*

[7] MIT BIH, Arrhythmia Database Directory. *Third edition, Harvard MIT Division of Health Sciences and technology, Biomed. Eng. Center. 1992.*

[8] T.Y. Lo, "A fast bandpass filter for ECG processing", *Proc. IEEE Eng. Med. Biol. Soc. 4th Annu. Int. Conf, pp: 20-21, 1982.*

[9] R. Wariar, "Integer coefficient bandpass filter for the simultaneous removal of baseline wander, 50 and 100 Hz interference from the ECG", *Med & Biol. Eng. & comput.,* 29,333-336, 1991.

[10] N.V. Thakor, "Design and analysis of quantized coefficient digital filters: application to biomedical signal processing with microprocessors", *Med & Biol. Eng. & Comput.* 25, 18-25, 1987.

[11] N.Belgacem, M.A. Chikh,F. Breksi-Reguig "Détection et identification des paramètres du signal ECG", *Séminaire National sur l'Automatique et les Signaux SNAS'02, Annaba : 27-28 octobre 2002.*

[12] C. Looney, "pattern recognition using neural" networks. *www.oup-usa.org: Oxford University press, 1997.*