A Neuro-Evolutionary Approach to Electrocardiographic Signal Classification

Antonia Azzini[1], Mauro Dragoni[2], and Andrea G. B. Tettamanzi[1]

[1] Università degli Studi di Milano
Dipartimento di Tecnologie dell'Informazione
via Bramante, 65 - 26013 Crema (CR) Italy
{antonia.azzini,andrea.tettamanzi}@unimi.it
[2] Fondazione Bruno Kessler (FBK-IRST)
Via Sommarive 18, Povo (Trento), Italy
dragoni@fbk.eu

Abstract. This work presents an evolutionary ANN classifier system as an heart beat classification algorithm suitable for implementation on the PhysioNet/Computing in Cardiology Challenge 2011 [7], whose aim is to develop an efficient algorithm able to run within a mobile phone, that can provide useful feedback in the process of acquiring a diagnostically useful 12-lead Electrocardiography (ECG) recording.
The method used in such a problem is to apply a very powerful natural computing analysis tool, namely evolutionary neural networks, based on the joint evolution of the topology and the connection weights together with a novel similarity-based crossover.
The work focuses on discerning between usable and unusable electrocardiograms tele-medically acquired from mobile embedded devices. A pre-processing algorithm based on the Discrete Fourier Trasform has been applied before the evolutionary approach in order to extract the ECG feature dataset in the frequency domain. Finally, a series of tests has been carried out in order to evaluate the performance and the accuracy of the classifier system for such a challenge.

Keywords: Signal Processing, Heartbeat Classification, Evolutionary Algorithms, Neural Networks

1 Introduction

In the last decades, cardiovascular diseases have represented one of the most important causes of death in the world [8] and the necessity of a trustworthy heart state evaluation is increasing. Electrocardiography (ECG) is one of the most useful and well-known methods for heart state evaluation. Indeed, ECG analysis is still one of the most common and robust solutions in the heart diseases diagnostic domain, also due to the fact that it is one of the simplest non-invasive diagnostic methods for various heart diseases [10].

In such a research field, one of the most important critical aspects regards the quality of such heart state evaluations, since, often, the lack of medically trained experts, working from the acquisition process to the discernment between usable and unusable medical information, increases the need of easy and efficient measuring devices, which can send measured data to a specialist. Furthermore, the volume of the data that have to be recorded is huge and, very often, the ECG records are non-stationary signals, and critical information may occur at random in the time scale. In this situation, the disease symptoms may not come across all the time, but would show up at certain irregular intervals during the day.

In this sense, the Physionet Challenge [7], on which this work focuses, aims at reducing, if not eliminating, all the fallacies that currently plague usable medical information tele-medically provided, by obtaining efficient measuring systems through smart phones.

In this challenge, several approaches were explored; in particular, in order to inform inexperienced user about the quality of measured ECGs, artificial-intelligence-based (AI-based) systems have been considered, by reducing the quantum of worst quality ECGs sent to a specialist, this contributing to a more effective use of her time.

Moody and colleagues [5] reported that some of the top competitors in this challenge employed a variety of techniques, using a wide range of features including entropy, higher order moments, intra-lead information, etc, while the classification methods also included Decision Trees, Support Vector Machines (SVMs), Fuzzy Logic, and heuristic rules.

An example of SVM-based approach is reported in [8], where the authors developed a decision support system based on an algorithm that combines simple rules in order to discard recordings of obviously low quality and a more sophisticated classification technique for improving quality of AI-based decision system for mobile phones, showing the fine tuning of sensitivity and specificity of detection. Another example has been also given in [9], where a rule-based classification method that mimics the SVM has been implemented, by using a modified version of a real time QRS-Complex algorithm and a T-Wave detection approach.

Anyway, according to [5], Artificial Neural Networks (ANNs) have been extensively employed in computer aided diagnosis because of their remarkable qualities: capacity of adapting to various problems, training from examples, and generalization capabilities with reduced noise effects. Also Jiang and colleague confirmed the usefulness of ANNs as heartbeat classifiers, emphasizing in particular evolvable ANNs, due to their ability to change the network structure and internal configurations as well as the parameters to cope with dynamic operating environments. In particular, the authors developed an evolutionary approach for the structure and weights optimization of block-based neural network (BbNN) models [18] for a personalized ECG heartbeat pattern classification.

We approach the heartbeat classification problems with another evolutionary algorithm for the joint structure and weights optimization of ANNs [4], which exploits an improved version of a novel similarity-based crossover operator [1], based on the conjunction of topology and connection weight optimization.

This paper is organized as follows: Section 2 briefly presents the problem, while a summary description of the evolutionary approach considered in this work is reported in Section 3. The results obtained from the experiments carried out are presented in Section 4, together with a discussion of the performances obtained. Finally, Section 5 provides some concluding remarks.

2 Problem Description

As previously reported, the ECG is a bio-electric signal that records the electrical activities of the heart. It provides helpful information about the functional aspects of the heart and cardiovascular system, and the state of cardiac health is generally reflected in the shape of ECG waveform, that is a critical information. For this reason, computer-based analysis and classification and automatic interpretation of the ECG signals can be very helpful to assure a continuous surveillance of the patients and to prepare the work of the cardiologist in the analysis of long recordings.

Moreover, as indicated by the main documentation of Physionet, according to the World Health Organization, cardiovascular diseases (CVD) are the number one cause of death worldwide. Of these deaths, 82% take place in low- and middle-income countries. Given their computing power and pervasiveness, the most important question is to check the possibility, for mobile phones, to aid in delivery of quality health care, particularly to rural populations distant from physicians with the expertise needed to diagnose CVD.

Advances in mobile phone technology have resulted in global availability of portable computing devices capable of performing many of the functions traditionally requiring desktop and larger computers. In addition to their technological features, mobile phones have a large cultural impact. They are user-friendly and are among the most efficient and most widely used means of communication. With the recent progress of mobile-platforms, and the increasing number of mobile phones, a solution to the problem can be the recording of ECGs by untrained professionals, and subsequently transmitting them to a human specialist.

The aim of the PhysioNet/Computing in Cardiology Challenge 2011 [6] is to develop an efficient algorithm able to run in near real-time within a mobile phone, that can provide useful feedback to a layperson in the process of acquiring a diagnostically useful ECG recording. In addition to the approaches already cited in Section 1, referring to such a challenge, Table 2 reports other solutions already presented in the literature, capable of quantifying the quality of the ECG looking at leads individually and combined, which can be implemented on a mobile-platform. As reported later, all such approaches are used to compare their results with those obtained in this work.

3 The Neuro-Evolutionary Algorithm

The overall algorithm is based on the evolution of a population of individuals, represented by Multilayer Perceptrons neural networks (MLPs), through a joint optimization of their structures and weights, here briefly summarized; a more complete and detailed description can be found in the literature [4]. In this work the algorithm uses the Scaled Conjugate Gradient method (SCG) [17] instead of the more traditional error back-propagation (BP) algorithm to decode a *genotype* into a *phenotype* NN, in order to speed up the convergence of such a conventional training algorithm. Accordingly, it is the genotype which undergoes the genetic operators and which reproduces itself, whereas the phenotype is used *only* for calculating the genotype's fitness. The rationale for this choice is that the alternative of applying SCG to the genotype as a kind of 'intelligent' mutation operator, would boost exploitation while impairing exploration, thus making the algorithm too prone to being trapped in local optima.

The population is initialized with different hidden layer sizes and different numbers of neurons for each individual according to two exponential distributions, in order to maintain diversity among all of them in the new population. Such dimensions are not bounded in advance, even though the fitness function may penalize large networks. The number of neurons in each hidden layer is constrained to be greater than or equal to the number of network outputs, in order to avoid hourglass structures, whose performance tends to be poor. Indeed, a layer with fewer neurons than the outputs destroys information which later cannot be recovered.

3.1 Evolutionary Process

The initial population is randomly created and the genetic operators are then applied to each network until the termination conditions are not satisfied.

At each generation, the first half of the population corresponds to the best $\lfloor n/2 \rfloor$ individuals selected by truncation from a population of size n, while the second half of the population is replaced by the offsprings generated through the crossover operator. Crossover is then applied to two individuals selected from the best half of the population (parents), with a probability parameter p_{cross}, defined by the user together with all the other genetic parameters, and maintained unchanged during the entire evolutionary process.

It is worth noting that the p_{cross} parameter refers to a 'desired' crossover probability, set at the beginning of the evolutionary process. However, the 'actual' probability during a run will usually be lower, because the application of the crossover operator is subject to the condition of similarity between the parents.

Elitism allows the survival of the best individual unchanged into the next generation and the solutions to get better over time. Following the commonly accepted practice of machine learning, the problem data is partitioned into training, validation and test sets, used, respectively for network training, to stop learning avoiding overfitting, and to test the generalization capabilities of a network. Then, the algorithm mutates the weights and the topology of the offsprings, trains the resulting network, calculates fitness on the validation set, and finally saves the best individual and statistics about the entire evolutionary process.

The application of the genetic operators to each network is described by the following pseudo-code:

1. Select from the population (of size n) $\lfloor n/2 \rfloor$ individuals by truncation and create a new population of size n with copies of the selected individuals.
2. For all individuals in the population:
   (a) Randomly choose two individuals as possible parents.
   (b) Check their local similarity and apply crossover according to the crossover probability.
   (c) Mutate the weights and the topology of the offspring according to the mutation probabilities.
   (d) Train the resulting network using the training set.
   (e) Calculate the fitness f on the validation set.
   (f) Save the individual with lowest f as the best-so-far individual if the f of the previously saved best-so-far individual is higher (worse).
3. Save statistics.

The `SimBa` crossover starts by looking for a 'local similarity' between two individuals selected from the population. If such a condition is satisfied the layers involved in the crossover operator are defined. The contribution of each neuron of the layer selected for the crossover is computed, and the neurons of each layer are reordered according to their contribution. Then, each neuron of the layer in the first selected individual is associated with the most 'similar' neuron of the layer in the other individual, and the neurons of the layer of the second individual are re-ranked by considering the associations with the neurons of the first one. Finally a cut-point is randomly selected and the neurons above the cut-point are swapped by generating the offspring of the selected individuals.

Weights mutation perturbs the weights of the neurons before performing any structural mutation and applying SCG to train the network. All the weights and the corresponding biases are updated by using variance matrices and evolutionary strategies applied to the synapses of each NN, in order to allow a control parameter, like mutation variance, to self-adapt rather than changing their values by some deterministic algorithms. Finally, the topology mutation is implemented with four types of mutation by considering neurons and layer addition and elimination. The addition and the elimination of a layer and the insertion of a neuron are applied with three independent probabilities, indicated as p^+_{layer}, p^-_{layer} and p^+_{neuron}, while the elimination of a neuron is carried out only if the contribution of that neuron is negligible with respect to the overall network output.

For each generation of the population, all the information of the best individual is saved.

As previously considered [3, 2], the evolutionary process adopts the convention that a lower fitness means a better NN, mapping the objective function into an error minimization problem. Therefore, the fitness used for evaluating each individual in the population is proportional to the mean square error (mse) and to the computational cost of the considered network. This latter term induces a selective pressure favoring individuals with reduced-dimension topologies.

The fitness function is calculated, after the training and the evaluation processes, by the Equation 1 and it is defined as a function of the confusion matrix M obtained by that individual:

$$f_{multiclass}(M) = N_{\text{outputs}} - \text{Trace}(M), \tag{1}$$

where N_{outputs} is the number of output neurons and $\text{Trace}(M)$ is the sum of the diagonal elements of the row-wise normalized confusion matrix, which represent the conditional probabilities of the predicted outputs given the actual ones.

4 Experiments and Results

The data used for the PhysioNet/CINC 2011 Challenge consist of 2,000 twelve-lead ECGs (I, II, III, aVR, aVF, aVL, V1, V2, V3, V4, V5, and V6), each 10 second long, with a standard diagnostic bandwidth defined in the range (0.05–100 Hz). The twelve leads are simultaneously recorded for a minimum of 10 seconds; each lead is sampled at 500 Hz with 16-bit resolution (i.e., 16 bits per sample).

The proposed approach has been evaluated by using the dataset provided by the challenge organizers. This dataset, described above in Section 2, is public and has been distributed in two different parts:

– Set A: this dataset has to be used to train the approach. It is composed of 998 instances provided with reference quality assessments;
– Set B: this dataset has to be used for testing the approach. It is composed of 500 instances and the reference quality assessments are not distributed to the participants. The reports generated by the approach have to be sent to the submission system in order to directly receive the results from the system used for the challenge.

We split the Set A in two parts: a training set composed of the 75% of the instances contained in the Set A, and a validation set, used to stop the training algorithm, composed of the remaining 25%. While the Set B is used as test set for the final evaluation of the approach.

Each instance of the dataset represents an ECG signal composed of 12 series (one for each lead) of 5,000 values representing the number of recordings performed for each lead. These data have been preprocessed in order to extract the features that we used to create the datasets given in input to the algorithm. We have applied to each lead the fast Fourier transform function in order to transform each lead to the frequency domain. After the transformation, we summed the 5,000 values by groups of 500 in order to obtain 10 features for each lead. Finally, The input attributes of all datasets have been rescaled, before being fed as inputs to the population of ANNs, through a Gaussian distribution with zero mean and standard deviation equal to 1.

The experiments have been carried out by setting the parameters of the algorithm to the values obtained from a first round of experiments aimed at identifying the best parameter setting. These parameter values are reported in Table 1. We performed 40 runs, with 40 generations and 60 individuals for each run, while the number of epochs used to train the neural network implemented in each individual has been set to 250.

Symbol	Meaning	Default Value
n	Population size	60
p^+_{layer}	Probability of inserting a hidden layer	0.05
p^-_{layer}	Probability of deleting a hidden layer	0.05
p^+_{neuron}	Probability of inserting a neuron in a hidden layer	0.05
p_{cross}	'Desired' probability of applying crossover	0.7
δ	Crossover similarity cut-off value	0.9
N_{in}	Number of network inputs	120
N_{out}	Number of network outputs	1
α	Cost of a neuron	2
β	Cost of a synapsis	4
λ	Desired tradeoff between network cost and accuracy	0.2
k	Constant for scaling cost and MSE in the same range	10^{-6}

Table 1. Parameters of the Algorithm.

The challenge has been organized in two different events: a closed event and an open one. While in the close event it is possible to develop the classification algorithm in any language, in the open event it is mandatory to develop the algorithm by using the Java language. For this reason, by considering that the proposed approach has been developed in Java too, we compared the obtained results with the results obtained by the other systems that participated to the challenge in the open event. It is important to highlight that we do not claim to obtain the best performance. Our goal was to show that, even if our system is trained with a training set that exploits very little information, the performance obtained by our approach does not lag too much behind the one obtained by the best state-of-the-art systems.

Table 2 shows the results obtained by the other participants compared with the results obtained by the proposed approach. Besides the comparison with the other approaches presented at the challenge, we have also compared our approach with the other following neuro-genetic approaches:

- Simple ANN with Conjugated Gradient: the classifiers are encoded with a population of ANNs, trained with the Conjugated Gradient method over

1,000,000 epochs. Also in this case, the networks are then evaluated over, respectively, the validation and the test sets, through the application of the mean square error.
– NeuroEvolution of Augmenting Topologies (NEAT) approach [19]: an evolutionary approach applied to neural network design that: (1) uses a crossover on different topologies, (2) protects structural innovation by using speciation, and (3) applies an incrementally growing from minimal network structures.
– Evolved ANN without crossover: the population of ANNs are evolved through the joint optimization of architecture and connection weights reported in this work, but in this case no crossover is implemented. The number of epochs corresponds to 250.

We have inserted both the best and the average performance obtained by the proposed approach. It is possible to observe that, if we consider the best performance, we obtained the second best accuracy; while the average accuracy, computed over the 40 runs, obtained the fourth performance. The robustness of the approach is also proved by observing the low value of the standard deviation that, in the performed experiments, was 0.011. With the italic font we show the performance obtained by the other approaches that we have used for classify the data in order to compare them with the approach proposed in this paper. The results demonstrated that the proposed approach outperforms the other ones. Indeed, the NEAT approach obtained only the seventh accuracy, while the other two approaches obtained respectively the eighth and the eleventh performance.

Participant	Score
Xiaopeng Zhao [11]	0.914
Proposed Approach (Best)	**0.902**
Benjamin Moody [12]	0.896
Proposed Approach (Average)	**0.892**
Lars Johannesen [13]	0.880
Philip Langley [14]	0.868
NEAT (Average)	*0.856*
Evolved ANN without crossover (Average)	*0.845*
Dieter Hayn [15]	0.834
Vclav Chudcek [16]	0.833
Simple ANN with Conjugated Gradient (Average)	*0.818*

Table 2. Results of the open event challenge.

Besides the evaluation on the test set, we performed also a ten-fold cross validation on the training set. We split the training set in ten fold F_i and we performed ten different set of 10 runs in order to observe which is the behavior of the algorithm when training, validation, and test data change. Table 3 shows the results of the ten-fold cross validation. By observing the results we can observe the robustness of the algorithm. In fact, the accuracies obtained by changing the folds used for training, validation, and test are very close; moreover, the standard deviation of the results is very low.

Training Set	Validation Set	Test Set	Average Accuracy	Standard Deviation
F1...F7	F8, F9	F10	0.8984	0.0035
F2...F8	F9, F10	F1	0.8988	0.0067
F3...F9	F10, F1	F2	0.9002	0.0075
F4...F10	F1, F2	F3	0.9022	0.0107
F5...F10, F1	F2, F3	F4	0.9040	0.0071
F6...F10, F1, F2	F3, F4	F5	0.9002	0.0029
F7...F10, F1...F3	F4, F5	F6	0.9002	0.0018
F8...F10, F1...F4	F5, F6	F7	0.8976	0.0054
F9, F10, F1...F5	F6, F7	F8	0.9032	0.0090
F10, F1...F6	F7, F8	F9	0.8986	0.0047

Table 3. Results of the ten-fold cross validation.

5 Conclusions

In this study, we have proposed an ECG classification scheme based on a neuro-evolutionary approach, based on the joint evolution of the topology and the connection weights together with a novel similarity-based crossover, to aid classification of ECG recordings. The signals were first preprocessed into the frequence domain by using a Fast Fourier Trasform algorithm, and then they were normalized through a gaussian distribution with 0 mean and standard deviation equal to 1. The present system was validated on real ECG records taken from the PhysioNet/Computing in Cardiology Challenge 2011.

A series of tests has been carried out in order to evaluate the capability of the neuro-evolutionary approach to discern between usable and unusable electrocardiograms tele-medically acquired from mobile embedded devices. The obtained results show an overall satisfactory accuracy and performances in comparison with other approaches carried out in this challenge and presented in the literature.

It is important to stress the fact that the proposed method was able to achieve top-ranking classification accuracy despite the use of a quite standard preprocessing step and a very small number of input features. No attempt was made yet to fine tune the signal pre-processing and the feature selection steps. On the other hand, it is well known that these two steps are often critical for the success of a signal classification methods. For this reason, we believe that the proposed neuro-evolutionary approach has a tremendous improvement potential.

Future works will involve the adoption of more sophisticated preprocessing techniques, by working, for example, on a multi-scales basis, where each scale represents a particular feature of the signal under study. Other ideas could regard the study and the implementation of feature selection algorithms in order to provide an optimized selection of the signal given as inputs to the neural networks.

References

1. Azzini, A., Tettamanzi, A., Dragoni, M.: SimBa-2: Improving a Novel Similarity-Based Crossover for the Evolution of Artificial Neural Networks. In: 11th International Conference on Intelligent Systems Design and Applications (ISDA 2011), pp. 374–379. IEEE (2011)
2. Azzini, A., Dragoni, M., Tettamanzi, A.: A novel similarity-based crossover for artificial neural network evolution. In: Parallel Problem Solving from Nature PPSN XI, Lecture Notes in Computer Science. vol. 6238, pp. 344–353. Springer (2010)
3. Azzini, A., Tettamanzi, A.: Evolving neural networks for static single-position automated trading. Journal of Artificial Evolution and Applications 2008(Article ID 184286), 1–17 (2008)
4. Azzini, A., Tettamanzi, A.: A new genetic approach for neural network design. In: Engineering Evolutionary Intelligent Systems. Studies in Computational Intelligence. vol. 82. Springer (2008)
5. Silva, K., Moody, G.B., Celi, L.: Improving the Quality of ECGs Collected Using Mobile Phones: The PhysioNet/Computing in Cardiology Challenge 2011. Contribution sent to the 38th Physionet Cardiology Challenge (2011).
6. PhysioNet: Research Resource for Complex Physiologic Signals, `http://www.physionet.org`
7. Moody, G.B.: Improving the quality of ECGs collected using mobile phones: The 12th Annual Physionet/Computing in Cardiology Challenge. Computing in Cardiology Challenge 38 (2011).
8. Data Driven Approach to ECG Signal Quality Assessment using Multistep SVM Classification. Contribution sent to the 38th Physionet Cardiology Challenge (2011).
9. Tat, T.H.C., Chen Xiang, C., Thiam, L.E.: Physionet Challenge 2011: Improving the Quality of Electrocardiography Data Collected Using Real Time QRS-Complex and T-Wave Detection. Contribution sent to the 38th Physionet Cardiology Challenge (2011).
10. S. Jokic, S. Krco, V. Delic, D. Sakac, Jokic, I., Lukic, Z.: An Efficient ECG Modeling for Heartbeat Classification. IEEE 10th Symposium on Neural Network Applications on Electrical Engineering, NEUREL 2010, Belgrade, Serbia, September, 23-25, 2010.
11. H. Xia, J. McBride, A. Sullivan, T. De Bock, J. Bains, D. Wortham, X. Zhao: A Multistage Computer Test Algorithm for Improving the Quality of ECGs. Contribution sent to the 38th Physionet Cardiology Challenge (2011).
12. B.E. Moody: A Rule-Based Method for ECG Quality Control. Contribution sent to the 38th Physionet Cardiology Challenge (2011).
13. L. Johannesen: Assessment of ECG Quality on an Android Platform. Contribution sent to the 38th Physionet Cardiology Challenge (2011).
14. P. Langley, L. Di Marco, S. King, C. Di Maria, W. Duan, M. Bojarnejad, K. Wang, D. Zheng, J. Allen, A. Murray: An Algorithm for Assessment of ECG Quality Acquired Via Mobile Telephone. Contribution sent to the 38th Physionet Cardiology Challenge (2011).

15. D. Hayn, B. Jammerbund, G. Schreier: Real-time Visualization of Signal Quality during Mobile ECG Recording. Contribution sent to the 38th Physionet Cardiology Challenge (2011).
16. V. Chudcek, L. Zach, J. Kulek, J. Spilka, L. Lhotsk: Simple Scoring System for ECG Signal Quality Assessment on Android Platform. Contribution sent to the 38th Physionet Cardiology Challenge (2011).
17. Hestenes, M., Stiefel, E.: Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standards 49(6) (1952).
18. W. Jiang, S.G. Kong. Block-Based Neural Networks for Personalized ECG Signal Classification. IEEE Transactions on Neural Networks, 18(6) (2007).
19. K. Stanley and R. Miikkulainen. Evolving neural networks through augmenting topologies. Evolutionary Computation, 10:99–127, 2002.