Classification of signals by means of genetic programming

Enrique Fernández-Blanco, Daniel Rivero, Marcos Gestal, Julián Dorado

University of A Coruña, Faculty of Computer Science, A Coruña, Spain

Abstract
This paper describes a new technique for signal classification by means of Genetic Programming (GP). The novelty of this technique is that no prior knowledge of the signals is needed to extract the features. Instead of it, GP is able to extract the most relevant features needed for classification. This technique has been applied for the solution of a well-known problem: the classification of EEG signals in epileptic and healthy patients. In this problem, signals obtained from EEG recordings must be correctly classified into their corresponding class. The aim is to show that the technique described here, with the automatic extraction of features, can return better results than the classical techniques based on manual extraction of features. For this purpose, a final comparison between the results obtained with this technique and other results found in the literature with the same database can be found. This comparison shows how this technique can improve the ones found.

1. Introduction

Signal processing and signal analysis have been a hot research topic over years. The main reason is that a huge part of real world problems produces signals which characterize them. So, any new knowledge about improving the processing can have a great soundness in other topics, such as medicine or engineering (Rabuñal et al. 2007; Ahsan et al. 2009).

Signal processing and signal analysis are a general topic which cover questions like signal characterization (Mallat and Hwang 1992), signal classification (Anderson et al. 1998), spectrum analysis (Lopes and Betrouni 2009), etc. The presented work can be framed under the epigraphs of automatic feature selection and signal classification.

This task is generally divided into three consecutive steps: pre-processing, feature extraction and classification. Pre-processing includes all the techniques used to improve the quality of the signals to be classified in order to extract more accurate features. At this step, processes like noise filtering, Blind Signal Separation (BSS) (Cardoso 1998) or independent component analysis (ICA) (Hyvarinen and Oja 2000) are used. This is quite important, because preprocessing can erase features that cannot be used in the following steps or it can improve or highlight features that could be mistaken with noise.

The next step is feature extraction, where different features are extracted from a signal in order to characterize it for the following classification. It is very important that the extracted features allow the separation of the classes in two (or more) regions of the problems' search space. In general, there are a lot of techniques that allow the feature extraction from different types of signals, after carrying out various types of analysis, such as entropy, frequency or wavelet analysis (Torrence and Compo 1998), Lyapunov exponents (Rosenblum et al. 1996), etc. From all these techniques, features that characterize the original signals more or less accurately can be extracted. The main problem of this step is that, usually, the features extracted from a signal depend on previous knowledge of the signals. Therefore, lacks in the knowledge can lead to having a worse set of features, and, at the end, to have lower classification accuracy. This problem can be treated by using an automatic feature extraction technique, which does not use previous knowledge in the extraction. The main objective of this work is to present a technique which is an alternative for this automatic feature extraction. It may be highlighted that automatic extraction can have as side effect the discovery of new knowledge, which is not dependent of the previous one.

Finally, the last step is the classification of the signal by using the previously extracted features. This classification is executed by one of the many different classification systems such as Artificial Neural Networks (ANNs) or Support Vector Machines (SVMs), which have been used in many works such as (Bazi and Melgani 2006; Zhan and Shen 2006; Rivero et al. 2011b). These techniques can generalize the knowledge presented in a dataset to determine the class which a new signal belongs to. It depends on the nature of the classifier used if prior information is required regarding to which class each signal belongs (supervised learning) or not (unsupervised learning), or if the classifying system needs prior training or not.

This paper present a new model for signal classification and automatic feature extraction based on Genetic Programming (GP). By using GP, resulting technique mixtures automatic extraction of the features and classification of the signal in one step. In order to illustrate the capacities of the proposed model, the authors have applied this new model to the classification of electroencephalogram signals regarding the epilepsy disease. Results show how this technique can improve those of other works over this topic with a more general technique.

On Sect. 2, it can be found a general state of the art, authors have considered to include a brief resume of GP (Sect. 2.1) and a deep summary of different works on automatic feature extraction and classification (Sect. 2.2) and EEG signals classification works (Sect. 2.3). Section 3 contains the whole description of the model and how it was tuned for EEG problem. The tests performed with this system and comparison with previous works is on Sect. 4. Finally, Conclusions of the work performed and future developments are in Sects. 5 and 6, respectively.

2. State of the art

2.1. Genetic Programming

Genetic Programming (GP) (Koza 1992) is an Evolutionary Computation technique, which has been successfully applied to many kinds of problems. As many Evolutionary Computation techniques, GP is inspired in the Theory of Evolution by imitating the evolutionary behavior of species. This technique uses a set of solutions called population which is composed by different solutions which are the individuals of the population. Each of these individuals is a tree which codifies the program to solve the problem.

Starting from a random initial population, this population is evolved through selection, mutation and crossover operators. These operators act like their counterparts in the natural evolution: in the population, the best individuals have a higher chance to reproduce themselves. By means of applying these operators, the population follows an iterative process which goes through different states, each of which is called a generation. As a result of this process, the population is expected to reach a generation that contains a good solution to the problem.

The ability of GP to build trees with a particular shape is the main advantage of this technique from the other Evolutionary Algorithms (EA) techniques. As GP allows the obtaining of trees, it is much more versatile than other techniques, like Genetic Algorithms, that have the restriction that they can only work with bit or floating point string. This codification does not allow, in this problem, the deep analysis of a signal by means of the combination of features extracted from it. On the other side, the use of a tree will allow the analysis of a signal with no limitations on the frequency bands used, or in the number of the frequency bands, and no limits on the combination of this information in order to classify the signal. This possibility has not been found in other EAs algorithms.

This technique has been applied in several problems of different fields with good results such as Civil Engineering, Medicine Applications, Signal Analysis, etc. (Rivero et al. 2005; Kishore et al. 2000; Hong et al. 2005; Dolinsky et al. 2007; Espejo et al. 2010).

2.2. Automatic feature extraction and classification

Feature extraction is usually performed through the analysis of the signal by means of any technique. From this analysis, and taking as a basis the knowledge that the expert has about the signals, several features are extracted. These features allow the classification of the signals in a more or less accurate way, depending on how accurately these features characterize the signals. However, usually the experts do not have a complete understanding and knowledge about the signals, and this leads to having features that do not give classification accuracy high enough. This problem can be overcome by means of using an automatic feature extraction technique, which allows the expert not to have previous knowledge about the signals. As the features are not extracted by the expert, but they are automatically extracted by the technique, they are expected to characterize the signals more accurately and therefore, the classification accuracy will be higher. It also makes possible a better classification and the discovery of new knowledge, not previously achieved by other studies. In this work, the features extracted are based on frequency analysis of the signals. Despite the large number of applications, in which signal classification methods are related, few methods for automatic frequency bands selection for classification have emerged.

To automatically select frequency bands, a possible approach consists of the analysis of different frequency bands, their feature extraction and subsequently the use of some technique for an automatic selection of whichever features are better for the classification task. An example of this type of system is described by Schröder et al. (2003), in which a Genetic Algorithm (GA) is used for the automatic selection of channels for EEG signal classification using a SVM. However, in this work, although the channels are selected, there is no selection of frequency bands in any of them. In another study, different features of signals are extracted, such as energy, fractal

dimension, or dominant frequency, and the authors have attempted to identify which features are the most significant for classification (Deriche and Al-ani 2001; Hsu et al. 2012).

A possible alternative could be to test of a great number of parameters in order to find the best accurate. For example, in (Lima and Coelho 2011) it is found a comparative of 26 wavelet kernel decomposition and the use of different types of SVM to classify them. The problem of this approach is the huge time that it requires and, usually, the results are not generally enough for another dataset.

In order to get an example of automatic selection of frequency bands, an important work was done by Dalponte. This work (Dalponte et al. 2007) is one in which an iterative strategy of exhaustive search is employed over EEG signals. In this work, all the possible ranges of time and frequency, between a minimum and a maximum given value over the EEG signals, were tested. However, the proposed algorithm returns a single frequency range from which the features are extracted, as the information contained may be at certain intervals within that range. Therefore, these intervals could give more information if they were known and returned by the algorithm in a finest level.

2.3. Epileptic EEG signal classification

EEG signal processing is a topic that has had a great impact from its beginning. Epileptic EEGs have been analyzed with many different tools and therefore their processing is very well documented (Mohseni et al. 2006).

One of the first techniques applied for this task was time and frequency analysis. Time-frequency features are also powerful tools for EEG signal analysis. This technique was used in some works, using the pseudo Wigner–Ville and the smoothed-pseudo Wigner–Ville distribution (Tzallas et al. 2007). For example, features were extracted and used as inputs to a feed-forward backpropagation neural network. In other works, an analysis in time and frequency domain is also provided, in which different features based on energy are extracted, features that are subsequently used also as inputs to a neural network (Rivero et al. 2009; Buteneers et al. 2011). Besides ANNs, other techniques have been used to classify the time-frequency features. For example, Polat and Günes (2007) proposed a frequency analysis using the Welch method, along with a decision tree.

In addition to time–frequency features, many different types of signal analyses have been performed over the signals, in order to extract features for characterize the signal. One of the most powerful tools for signal and image processing is the wavelet transform (Mohseni et al. 2006); (Addison 2002). For example, Subasi and Gursoy (2010), analyses the EEG signals by means of the discrete wavelet transform, and then he extract the statistical features of the sub bands by using the principal component analysis (PCA), independent component analysis (ICA) and linear discriminant analysis (LDA). Those statistical features are then used as inputs of an SVM, which performs the final classification. In a different work, features based on the Relative Wavelet Energy have been extracted and used as input of an ANN (Guo et al. 2009). Another alternative is to use the wavelet features with a neuro-fuzzy system like in (Guler and Ubeyli 2005).

Other different analysis of the signal is done by the extraction of entropy-based features. Different entropy estimators have been applied to quantify the complexity of a signal. This led to the extraction of different features that were used as inputs of an ANFIS classifier (Kannathal et al. 2005b).

Another approach refers to the study of EEG signals as chaotic systems. From this point of view, Lyapunov exponents are extracted from time series using Jacobi matrices (Abarbanel et al. 1991). From these exponents, different features can be extracted, which are used for the classification (Übeyli 2009). In other works, methods based on a non-linear dynamics theory were used to extract features based on the fractal dimension, along with a SVM classifier (Schneider et al. 2009).

3 Model description

On this paper, the authors propose the use of Genetic Programming (GP) in order to solve the problems described on Sect. 4.1. In order to do this, GP's typing property (Montana 1995) is used to generate trees with particular shapes and properties, which will be used to solve the problem.

The system described on this section can be used to classify the signals contained on the dataset. The main feature of this system is the automatic extraction of the frequency information to classify the signal without any previously knowledge about the signal. In order to do this, each one of the signals was previously analyzed by means of their Fast Fourier Transform (FFT) into frequency and the Power Spectrum Density (PSD).

Usually, before this analysis, the signals undergo a pre-processing stage that includes some steps such as noise reduction. In this case, no pre-processing is used, because it usually leads to losing some information. Noise reduction usually implies erasing information in selected (and wide) frequency bands. In this case, as frequency analysis is done, the elimination of the information in those frequency bands means that that information could not be used as input to the system. Although it could be thought that this information useless, in this paper it is preferred not to delete it, because the feature extraction process is completely automated, and this technique could find useful information in those frequency bands. In this case, if that information is actually useful, the accuracy of the system would be higher. If noise reduction was applied, no useful information could be extracted from these bands, and, if the deleted information was useful, the results would be worse.

With this information from the frequency analysis, a GP individual is able to classify a signal by using the information on time domain or frequency domain (FFT and PSD, in absolute value). In order to perform this classification, the GP individual will have a set of internal nodes of the tree, which will be to recover information from these domains by applying a specific function. These internal nodes of the tree will have two children which will designate the parameters of the function as real values, and they also return real values as result of their evaluation. These nodes are the following:

- MeanTime and StdTime. Their two parameters (children) are indexes of two samples in time domain. With these two indexes, these nodes perform the average and the standard deviation of the chunk of the signal between those indexes.
- MeanFFT and StdFFT. Their two parameters (children) set two frequencies. With these two frequencies, these nodes perform the average and the standard deviation of the signal's FFT between these two frequencies.
- MeanPSD and StdPSD. Again, the two parameters (children) set two frequencies. With them, these nodes perform the average and the standard deviation of the signal's PSD between these two frequencies.

An important restriction added to the GP trees is that none of them can have any of these six nodes inside another. This means that the sub-trees of these nodes will designate constant values, and these values will be the same for each of the signals analyzed by its tree.

As it was previously mentioned, these nodes take as arguments the result of the evaluation of their children sub-trees. The sub-trees have a REAL type, and these nodes also return a REAL type. In order to allow the sub-trees to generate real constant values, it is necessary to add to the terminal and function sets some nodes that perform arithmetic operations, as well as random constants.

Moreover, the output of the GP individual must be a class value, since the individual must classify the signal in one of the several classes. In order to perform this, the trees are asked to be of a certain type: CLASS, and the following nodes are added to the terminal and function sets:

- "If-then-else": node of CLASS type. It has three children, the first with BOOLEAN type and the other two with CLASS type. If the result of the evaluation of the first child is true, then the output of this node will be the result of the evaluation of the second child. Otherwise, the output will be the result of the evaluation of the third child.
- As boolean operations are needed, relational nodes are included. These nodes perform comparison operations between real values. Therefore, they have two children, with REAL type, and they return a BOOLEAN type value.
- A terminal node for each class in which the signals can be classifies is added to the terminal set. For instance, if the problem to be solved is A–B, i.e., to separate class A from B, nodes "A" and "B" are added to the terminal set. If the problem to solve is A–B–C, then the nodes "A", "B" and "C" are added to the terminal set. All of these nodes have a CLASS type.

Table 1 shows a summary of the terminal and functions sets used in this work to build the trees. In this Table, symbol "%" stands for the protected division, which returns a value of 1 if its second argument is 0.

Table 1. Summary of terminal and functions used by the GP

Name	Type	Num of children	Type of children
Function set			
MeanTime			
StdTime			
MeanFFT			
StdFFT	REAL	2	REAL, REAL
MeanPSD			
StdPSD			
+, −, *, %	REAL	2	REAL, REAL
<, >, =, ≠	BOOLEAN	2	REAL, REAL
If–then–else	CLASS	3	BOOLEAN, CLASS, CLASS
Terminal set			
[−1, 1]	REAL	−	−
Class1, class2, etc.	CLASS	−	−

Figure 1 shows an example of a possible tree built with the operations of Table 1 to solve the ZONF–S problem.
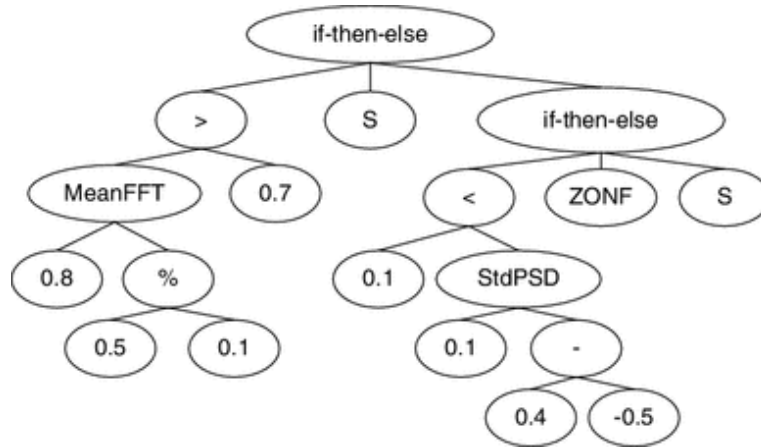
Fig. 1. Example of GP individual as tree notation

Figure 1 tree represents the following pseudo-code algorithm:

```
if (mean of the FFT values between 0.8Hz and 5Hz > 0.7)
then output class = S
else
        if (0.1 < standard deviation of the PSD values between 0.1Hz and 0.9Hz)
        then output class = ZONF
        else output class = S
```

To set the fitness of the different GP trees, each one of them is evaluated with every signal on the dataset. The fitness value of each individual will be the average of the misclassification over the signals on that database made by the tree. The general function is on Eq. 1, where o_i is the output of the classification system, s_j is signal desired classification and *#patterns* is the number of patterns used in the tests.

$$\text{Fitness} = \sum_{i=1}^{\#\text{patterns}} (e_i) \,/ \#\text{patterns}, e_i = \begin{cases} 0 & o_i = s_i \\ 1 & o_i^1 s_i \end{cases} \tag{1}$$

4. Results

This section describes the different tests that have been performed with the proposed method and a comparison with the results achieved by other techniques on the same data.

4.1. Problem description

The practical application of this paper refers to the classification of EEG signals. The dataset used in this work is composed by signals from healthy people and epileptic patients. Epilepsy is characterized by the occurrence of seizures in the EEG signal. The database used in this article contains several EEG recordings, is available publicly, and has been described in previous works (Andrzejak et al. 2001).

This database is composed by five datasets (named Z, O, N, F and S). Each one of this datasets contains 100 single-channel EEG signals. Each of these signals lasts 23.6 s, with a sampling frequency of 173.61 Hz. Therefore, the total length of each segment is of 4,097 samples. These segments were selected and cut out from continuous multi-channel EEG recordings after visual inspection for artifacts (e.g., due to muscle activity or eye movements). In addition, the segments had to meet a criterion of stationarity. An example of each signal is on Fig. 2.
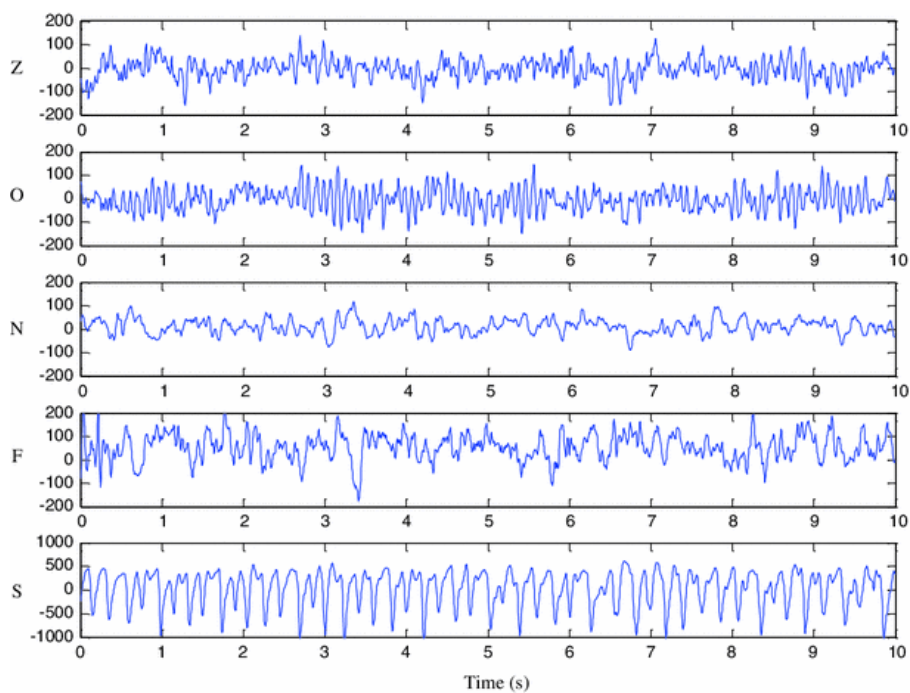
Fig. 2. Example of EGG signals kinds

Datasets named as Z and O, consisted of segments taken from surface EEG recordings that were carried out on five healthy volunteers with eyes open (Z) and closed (O). On the other hand, datasets N, F and S originated from an EEG archive of presurgical diagnosis. Samples in dataset N were recorded from the epileptogenic zone, and those in dataset F from the hippocampal formation of the opposite hemisphere of the brain. While sets N and F contained only activity measured during seizure free intervals from epileptic patients, set S only contained seizure activity. Here, segments were selected from all the recording sites exhibiting ictal activity. Reference data of the different signals are in Table 2.

Table 2. Principal indexes of the kinds signals

Set	Mean	Std.	Max.	Min.
Z	−6.26	40.73	294	−288
O	−12.51	61.11	360	−424
N	8.88	50.83	623	−412
F	−6.20	65.62	2,047	−1,147
S	−4.75	306.61	2,047	−1,885

The aim of this work is to perform a classification in several classes, depending on the data sets taken into account. For example, those segments that show seizure activity (S) and those which do not show it (non seizure, Z, O, N, F), either belonging to an epileptic patient or to a healthy person. This specific problem (ZONF-S) was chosen because it is closer to the clinical applications. Other possible problems studied in this work are Z–S, Z–F–S or ZO–NF–S, having different data sets size. The authors have compiled results from different works of the state of the art and here the system faces everyone of these problems that are showed on Table 3.

Table 3. Problem characterization

| Problem	Number of classes	Patterns
Z-S	2	200
ZONF-S	2	500
Z–F–S	3	300
ZO–NF–S	3	500

4.2. Description of the experiments

It has to be highlighted that all of the tests have tried to be performed in the same conditions for each technique. Previously works split the data set into 50 % of the data for training and 50 % for testing. In this work, this condition has been kept in order to compare the results with other works. It has to be highlighted that partitions of previously works are not available, so in order to reduce the influence of the randomness of the evolutionary research technique each test has been carried out 50 times. Therefore, the results presented are the average of those 50 tests.

The data used in the tests, as it was commented in Sect. 4.1, are from a public accessible signals database (Andrzejak et al. 2001). The dataset are 100 signals of EEGs with 4,097 samples for each one of the signals. These samples were normalized between 0 and 1.

The GP system was tuned up with the following values for its configuration, these values were empirically selected after some previously tests:

- Number generations: 1,000.
- Population size: 1,000.
- Crossover percent: 95 %.
- Mutation probability: 4 %.
- Selection operation: tournament of 2 individual.
- Maximum tree depth: 9.
- Creation algorithm: Ramped half and half.

It has to be mention that GP uses an elitism strategy on the usual crossover and mutation operations. This means that the best individual of the population is copied to the new generation without any modifications to ensure that the population doesn't achieve worse results than the previous. This tries to keep the best individual over the generations which, in other case, could be lost due to the use of evolutionary operators. A schema of the overall behavior of the system for one of the test can be seen on Fig. 3. This process is repeated fifty times in order to calculate the average performance of the system.
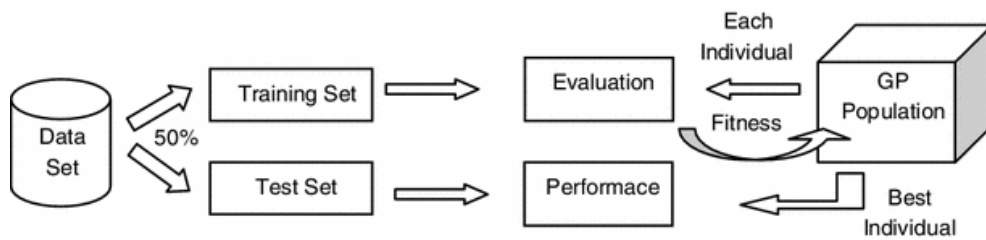
Fig. 3. Overall Ttraining and testing for one individual

4.3. Experimental results

Table 4 shows the overall results for the different classification problems with the proposed technique. This Table shows the average results for 50 independent trials performed for each problem, and, for each one, the accuracy and standard deviations in training and test can be seen. It also shows the overall accuracy and the percent of runs that returned a 100 % accuracy in test.

Table 4 Resume of the result for 50 executions for each problem

	Training		Test			
	Accuracy (%)	Std.	Accuracy (%)	Std.	% 100 % in test	Overall accuracy (%)
Z-S	100	0	100	0	100	100
ZONF-S	100	0	99.38	0.481	22	99.69
Z–F–S	100	0	98.51	1.12	14	99.25
ZO–NF–S	99.99	0.06	97.86	1.06	0	98.93

As this Table shows, the overall accuracy of the method is higher than 98.9 % for every problem. In fact, the results of the test are remarkable. For example, the Z–S problem which accuracy for tests is 100 %, so the fifty executions of the system have achieved a solution for this problem. The rest of the problems achieve average accuracies over the 97 %.

4.4 Comparative results

Results obtained herein were compared with the ones reported by other works. This comparison is motivated to evaluate the proposed system.

In order to perform a fair comparison, the experiments have to be done with the same conditions as in the works used for the comparison. In this case, the conditions are almost identical, because the database is the same. The only difference is the division of the database into the training and test sets. However, this division is not provided in any of the reference papers with the reported results. In this work, the same methodology was used: split the database into training and test a great number of times and, at the end, report the average of the results obtained. This average result can be compared with other results obtained in other works, and it is very usual to report the average results published in other works as a basis for the comparison. This is the reason why in this work the comparison is based only on the average of the different executions (shown in Table 4). This comparison is shown on Table 5.

Table 5 Comparative results with the state of the art

Problem	Reference	Method	Accuracy
Z–S	Nigam and Graupe (2004)	Non-lineal pre-processing filter-Diagnostic neural network	97.2
	Srinivasan et al. (2005)	Time–frequency domain features—recurrent neural network	99.6
	Kannathal et al. (2005b)	Entropy measures—adaptive neuro-fuzzy inference system	92.22
	Kannathal et al. (2005a)	Chaotic measures—surrogate data analysis	~90
	Polat and Günes (2007)	Fast Fourier transform—decision tree	98.72
	Subasi (2007)	Discrete wavelet transform—mixture of expert model	95
	Tzallas et al. (2009)	Time frequency analysis, artificial neural network	100
	This work		**100**
ZONF-S	Tzallas et al. (2007)	Time frequency analysis, artificial neural network	97.73
	Rivero et al. (2011a)	Time frequency analysis, kNN	98.40
	Rivero et al. (2011b)	Time frequency analysis, RNN	98.60
	This work		**99.38**
Z–F–S	Guler et al. (2005)	Lyapunov exponents—recurrent neural network	96.79
	Sadati et al. (2006)	Discrete wavelet transform—adaptive neural fuzzy network	85.9
	Tzallas et al. (2009)	Time frequency analysis, artificial neural network	100
	This work		**98.51**
ZO–NF–S	Tzallas et al. (2007)	Time frequency analysis, artificial neural network	97.72
	This work		**97.86**

Bold values indicate the best results

In Table 5, it is clear that the proposed system achieves better results than the rest of the works with the exception of problem Z–F–S. Results of (Tzallas et al. 2009) is better than the system proposed here but it may be highlighted that in this work the author uses the overall accuracy performance instead of test results as is done in this work. In the rest of the problems, the proposed system achieves the better results among all the works.

5. Conclusions

Once the test results have been presented, some conclusions can be extracted from this work. The main conclusion is that GP can be used for the classification of signals. In the particular case of the application described here, this technique returned the best results in almost all of the problems.

It may be also highlighted that the other works use a combination of techniques. This work joins the automatic feature extraction and the classification, oppositely to other techniques, which divide this process into two different steps with two different techniques. This technique does not need a previous feature extraction process, since the features (in this case, frequency bands) are automatically selected for each GP individual by the tree nodes. This is an advantage, since no prior knowledge is needed to classify the signals.

6. Future works

This work has presented the application of a new system for classification signals, and it has been applied to the classification of epileptic EEG signals. The proposed system could also be applied in other environments, involving EEG signals or other king of signals.

Also, in this work time and frequency domains are studied. However, the GP model could be adapted to work in other domains, such as wavelets, to allow GP to extract new features that could improve the classification.

References

Abarbanel HDI, Brown R, Kennel MB (1991) Lyapunov exponents in chaotic systems: their importance and their evaluation using observed data. Int J Mod Phys 5(9):1347–1375. doi: 10.1142/S021797929100064X

Addison PS (2002) The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance. Institute of Physics Publishing, Bristol

Ahsan MR, Ibrahimy MI, Khalifa OO (2009) EMG signal classification for human computer interaction: a review. Eur J Sci Res 33(3):480–501

Anderson CW, Stolz EA, Shamsunder S (1998) Multivariate autoregressive models for classification of spontaneous electroencephalographic signals during mental tasks. IEEE Trans Biomed Eng 45(3):277–286. doi: 10.1109/10.661153

Andrzejak RG, Lehnertz K, Rieke C, Mormann F, David P, Elger CE (2001) Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. Phys Rev E Stat Nonlin Soft Matter Phys 64 (6) doi: 061907

Bazi Y, Melgani F (2006) Toward an optimal SVM classification system for hyperspectral remote sensing images. IEEE Trans Geosci Remote Sens 44(11):3374–3385. doi: 10.1109/tgrs.2006.880628

Buteneers P, Verstraeten D, van Mierlo P, Wyckhuys T, Stroobandt D, Raedt R, Hallez H, Schrauwen B (2011) Automatic detection of epileptic seizures on the intra-cranial electroencephalogram of rats using reservoir computing. Artif Intell Med 53(3):215–223. doi: 10.1016/j.artmed.2011.08.006

Cardoso JF (1998) Blind signal separation: statistical principles. Proc IEEE 86(10):2009–2025. doi: 10.1109/5.720250

Dalponte M, Bovolo F, Bruzzone L (2007) Automatic selection of frequency and time intervals for classification of EEG signals. Electron Lett 43(25):1406–1408. doi: 10.1049/el:20072428

Deriche M, Al-ani A (2001) A new algorithm for EEG feature selection using mutual information. In: IEEE International Conference of the Acoustics Speech and Signal Processing 2001, pp 1057–1060. doi: 10.1109/ICASSP.2001.941101

Dolinsky JU, Jenkinson ID, Colquhoun GJ (2007) Application of Genetic Programming to the calibration of industrial robots. Comput Ind 58(3):255–264. doi: 10.1016/j.compind.2006.06.003

Espejo PG, Ventura S, Herrera F (2010) A survey on the application of genetic programming to classification. Systems, man, and cybernetics, Part C: applications and reviews. IEEE Transactions on 40 (2):121–144. doi: 10.1109/TSMCC.2009.2033566

Guler I, Ubeyli ED (2005) Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients. J Neurosci Methods 148(2):113–121. doi: 10.1016/j.jneumeth.2005.04.01

Guler NF, Ubeylib I, Guler ED, Guler I (2005) Recurrent neural networks employing Lyapunov exponents for EEG signals classification. Expert Syst Appl 29:506–514. doi: 10.1016/j.eswa.2005.04.011

Guo L, Rivero D, Seoane JA, Pazos A Classification of EEG signals using relative wavelet energy and artificial neural networks. In: Proceedings of the first ACM/SIGEVO Summit on genetic and evolutionary computation, Shanghai, China, 2009. pp 177–184. doi: 10.1145/1543834.1543860

Hong G, Jack LB, Nandi AK (2005) Feature generation using genetic programming with application to fault classification. In: IEEE Transactions on Systems, Man and Cybernetics, Part B: cybernetics 35 (1):89–99

Hsu WY, Lin CH, Hsu HJ, Chen PH, Chen IR (2012) Wavelet-based envelope features with automatic EOG artifact removal: application to single-trial EEG data. Expert Syst Appl 39(3):2743–2749. doi: 10.1016/j.eswa.2011.08.132

Hyvarinen A, Oja E (2000) Independent component analysis: algorithms and applications. Neural Netw 13(4–5):411–430. doi: 10.1016/s0893-6080(00)00026-5

Kannathal N, Acharya UR, Lim CM, Sadasivan PK (2005a) Characterization of EEG—a comparative study. Comput Methods Prog Biomed 80(1):17–23. doi: 10.1016/j.cmpb.2005.06.005

Kannathal N, Choob ML, Acharyab UR, Sadasivana PK (2005b) Entropies for the detection of epilepsy in EEG. Comput Methods Programs Biomed 80(3):187–194. doi: 10.1016/j.cmpb.2005.06.01

Kishore JK, Patnaik LM, Mani V, Agrawal VK (2000) Application of genetic programming for multi category pattern classification. IEEE Trans Evol Comput 4(3):242–258. doi: 10.1109/4235.873235

Koza J (1992) Genetic programming: on the programming of computers by means of natural selection. The MIT Press, Cambridge

Lima CAM, Coelho ALV (2011) Kernel machines for epilepsy diagnosis via EEG signal classification: a comparative study. Artif Intell Med 53(2):83–95. doi: 10.1016/j.artmed.2011.07.003

Lopes R, Betrouni N (2009) Fractal and multifractal analysis: a review. Med Image Anal 13(4):634–649. doi: 10.1016/j.media.2009.05.003

Mallat S, Hwang WL (1992) Singularity detection and processing with wavelets. IEEE Trans Infor Theory 38(2):617–643. doi: 10.1109/18.119727

Mohseni HR, Maghsoudi A, Shamsollahi B Seizure Detection in EEG signals: a comparison of different approaches. In: Conference of the IEEE Engineering in Medicine and Biology Society 2006, pp 6724–6727. doi: 10.1109/IEMBS.2006.260931

Montana DJ (1995) Strongly typed genetic programming. Evol Comput 3(2):199–230. doi: 10.1162/evco.1995.3.2.199 CrossRefGoogle Scholar

Nigam VP, Graupe D (2004) A neural-network-based detection of epilepsy. Neurol Res 26(1):55–60. doi: 10.1179/016164104773026534 CrossRefGoogle Scholar

Polat K, Günes S (2007) Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Appl Math Comput 32(2):625–631. doi: 10.1016/j.amc.2006.09.022

Rabuñal JR, Puertas J, Suarez J, Rivero D (2007) Determination of the unit hydrograph of a typical urban basin using Genetic Programming and artificial neural networks. Hydrol Process 21:476–485. doi: 10.1002/hyp.6250

Rivero D, Rabuñal JR, Dorado J, Pazos A (2005) Time series forecast with anticipation using Genetic Programming. Lect Notes Comput Sci 3512:968–975. doi: 10.1007/11494669_119

Rivero D, Dorado J, Rabuñal J, Pazos A (2009) Evolving simple feed-forward and recurrent ANN's for signal classification: A comparison. IEEE-INNS-ENNS International Joint Conference on Neural Networks, pp 2685–2692.doi: 10.1109/IJCNN.2009.5178621

Rivero D, Fernandez-Blanco E, Dorado J, Pazos A (2011a) A new signal classification technique by means of Genetic Algorithms and kNN. IEEE Congress on Evolutionary Computation (CEC), pp 581–586. doi: 10.1109/CEC.2011.5949671

Rivero D, Fernandez-Blanco E, Dorado J, Pazos A (2011b) Using recurrent ANNs for the detection of epileptic seizures in EEG signals. IEEE Congress on Evolutionary Computation (CEC), pp 587–592. doi: 10.1109/CEC.2011.5949672

Rosenblum MG, Pikovsky AS, Kurths J (1996) Phase synchronization of chaotic oscillators. Phys Rev Lett 76(11):1804–1807. doi: 10.1103/PhysRevLett.76.1804

Sadati N, Mohseni HR, Maghsoudi A (2006) Epileptic Seizure Detection using neural fuzzy networks. In: IEEE International Conference on Fuzzy Systems, pp 596–600 doi: 10.1109/FUZZY.2006.1681772

Schneider M, Mustaro PN Lima CAM (2009) Automatic recognition of epileptic seizure in EEG via support vector machine and dimension fractal. In: Proceedings of the 2009 international joint conference on Neural Networks, pp 2841–2845. doi: 10.1109/IJCNN.2009.5179059

Schröder M, Bogdan M, Rosenstiel W, Hinterberger T, Birbaumer N (2003) Automated EEG feature selection for brain computer interfaces. In: Proceedings of the 1st International IEEE EMBS Conference on Neural Engineering, Capri Island, Italy, pp 626–629. doi: 10.1109/CNE.2003.1196906

Srinivasan V, Eswaran C, Sriraam N (2005) Artificial neural network based epileptic detection using time-domain and frequency-domain features. J Med Syst 29(6):647–660. doi: 10.1007/s10916-005-6133-1

Subasi A (2007) EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Syst Appl 32(4):1084–1093. doi: 10.1016/j.eswa.2006.02.005

Subasi A, Gursoy MI (2010) EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst Appl 37:8659–8666. doi: 10.1016/j.eswa.2010.06.065

Torrence C, Compo GP (1998) A practical guide to wavelet analysis. Bull Am Meteorol Soc 79(1):61–78. doi: 10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2 CrossRefGoogle Scholar

Tzallas AT, Tsipouras MG, Fotiadis DI (2007) Automatic seizure detection based on time-frequency analysis and artificial neural networks. Comput Intell Neurosci 7(3):1–13. doi: 10.1155/2007/80510 CrossRefGoogle Scholar

Tzallas AT, Tsipouras MG, Fotiadis DI (2009) Epileptic seizure detection in EEGs using time–frequency analysis. IEEE Trans Infor Technol Biomed 13(5):703–710. doi: 10.1109/TITB.2009.2017939 CrossRefGoogle Scholar

Übeyli ED (2009) Lyapunov exponents/probabilistic neural networks for analysis of EEG signals. Expert Syst Appl 37(2):985–992. doi: 10.1016/j.eswa.2009.05.078

Zhan YQ, Shen DG (2006) Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method. IEEE Trans Med Imaging 25(3):256–272. doi: 10.1109/tmi.2005.862744