Epileptic Seizure Detection Using Genetically Programmed Artificial Features

Hiram Firpi*, *Member, IEEE*, Erik D. Goodman, and Javier Echauz, *Member, IEEE*

***Abstract*—Patient-specific epilepsy seizure detectors were designed based on the genetic programming artificial features algorithm, a general-purpose, methodic algorithm comprised by a genetic programming module and a *k*-nearest neighbor classifier to create synthetic features. Artificial features are an extension to conventional features, characterized by being computer-coded and may not have a known physical meaning. In this paper, artificial features are constructed from the reconstructed state-space trajectories of the intracranial EEG signals intended to reveal patterns indicative of epileptic seizure onset. The algorithm was evaluated in seven patients and validation experiments were carried out using 730.6 hr of EEG recordings. The results with the artificial features compare favorably with previous benchmark work that used a handcrafted feature. Among other results, 88 out of 92 seizures were detected yielding a low false negative rate of 4.35%.**

***Index Terms*—Epilepsy, feature extraction, genetic programming, seizure detection, state-space reconstruction.**

I. Introduction

APPROXIMATELY one percent of the world is affected by epilepsy, a bewildering neurological disorder that may cause brief electrical disturbances in the brain, producing a change in the sensation, awareness, and behavior and which is typified by recurrent seizures. Twenty-five percent of these cases cannot be fully controlled using current therapies, medications, or surgical treatments. For these people, a stimulation device would be of great benefic. From a generalistic point of view, an implantable device would be compounded by a detector component and a control/stimulation component. In this paper, our efforts are devoted to present a new tool for the detection of epileptic seizures.

The introduction of the electroencephalogram (EEG) has brought cues characterizing the dynamics of the epilepsy. Although detection of epileptic seizures is easier than the prediction problem, is still far from trivial. There are many anomalies that occur naturally in the EEG signals that trigger the detectors and declare that a seizure is occurring when actually it is not. False starts (i.e., high-frequency, low amplitude events similar to those occurring at the beginning of a seizure), delta trains, and spike-and-wave discharges, all of them lead to false alarms and, thus, would medicate the patient unnecessarily. Additionally, seizures must be detected as soon as possible so the control/stimulation can be delivered immediately and control the seizures without further consequences. Therefore, there is a tradeoff between the number of false alarms and the number of false negatives (when the detector says a seizure is going to occur and it does not), and consequently affecting how early a seizure can be detected.

Several algorithms have been proposed to detect epileptic seizure onsets [1], [5], [6], [10], [15], [18], [20], [21], [24], among others in the search for an accurate detector. Some approaches that apply digital signal processing or filter theory have acceptable performance. These algorithms rely on one or more features—informative measures or attributes eluded from raw data—to decide whether a seizure is occurring or not. However, to extract the relevant information that can facilitate such detection, features are calculated using conventional techniques and methodologies that are time-consuming, tedious, and trial-and-error processes requiring a great deal of effort from researchers. All of these conventional techniques rely on gathering a feature or a set of features conceived by knowledge of a feature formula or algorithm that may have been obtained from intuition, tradition, the physics of the problem, or analogies to problems in other fields. Because of this, we wonder if there is any underlying pattern or patterns that are being ignored by these traditional methods. The relevancy provided by the features is limited by the attributes measured for such features.

In pursuing a methodology that surpasses the aforementioned limitations of the conventional features, we develop an algorithm that systematically and automatically can find or generate artificial features or attributes starting from raw data—in this case, from the IEEG signals. Artificial features are defined by features that are computer-designed, learned, inductive, optimized, and data-driven. These are desirable properties on features because they are meant to pay attention, capture, extract, or uncover any underlying, useful pattern that might be suppressed by the conventional features. We will see that unlike most approaches, this work is not based solely on tuning a few parameters of fixed terms to find the "best" detector for a patient, but we are using an algorithm that also changes the structure of the equations, giving us enough flexibility to design an "optimal" (or at least highly tuned) feature that is sensitive to the characteristics or patterns of a given patient.

Section II explains the methodology on how to reconstruct the state-space trajectory from the IEEG signal. It also explains the

Manuscript received July 31, 2005; revised June 24, 2006. *Asterisk indicates corresponding author*.

*H. Firpi is with the Center for Computational Biology & Bioinformatics, Indiana University-Perdue University, 410 W. 10th Street, Suite 5000, Indianapolis, IN 46202 USA (e-mail: hfirpi@ieee.org).

E. D. Goodman is with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824-1226 USA (e-mail: goodman@egr.msu.edu).

J. Echauz is with the BioQuantix Corp., Atlanta, GA 30363 USA (e-mail: echauz@ieee.org).

Color versions of Figs. 2 and 5–7 are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2006.886936

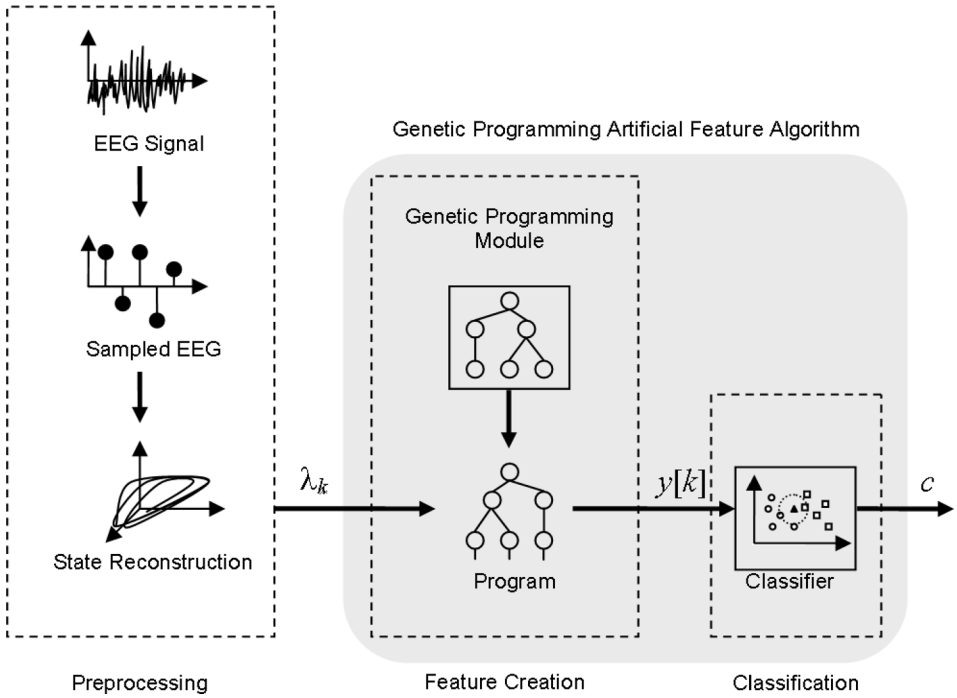


Fig. 1. State-reconstruction (preprocessing) and the components of the genetic programmed artificial algorithm (feature creation and classification, grayed square).

components that constitute the genetic programming artificial features (GPAF) algorithm. Section III shows the results of the experiments on data from seven patients. Sections IV and V present discussions and conclusions, respectively.

II. Methodology

We present a methodology in which we attempt to capture one or more deterministic components of the dynamics of IEEG signals by means of delay-embedding (i.e., a preprocessing stage) in a stream of sliding windows. First, the state-space trajectory of the IEEG signal is reconstructed using the standard delay-embedding scheme. Later, these reconstructed states are input into a genetic programming algorithm, which attempts to find one or more patterns giving the best discrimination between baseline data and ictal data (i.e., seizure data), in the sense of a minimum-error-risk objective function. A universal classifier then performs the detection (categorization) task. Fig. 1 depicts a diagram with all components of the algorithm. The next subsections expound on the above summary.

A. EEG Data

The anonymized intracranial electroencephalograph (IEEG) data used for these experiments were obtained from Emory University Hospital at Atlanta, GA, USA. The IEEG data were recorded from diagnosed mesial temporal lobe epileptic patients—strongly associated with complex partial seizures, the most common type of seizure, being present in 40% of all epilepsy cases reported [2]—undergoing a presurgical evaluation. To record the IEEG data, patients were implanted with bipolar amygdalo-hippocampal depth electrodes, subdural strip electrodes. The number of electrodes varied from 20 to 32. Fig. 2 shows the electrode contact arrangement. Additionally, patients were simultaneously videotaped during their hospital stays, which varied in duration from 4 to 11 days. The IEEG signals were recorded at 200 Hz, with 12-bit resolution. To digitally derive the bipolar IEEG data used for the experiments, the adjacent pair of electrode contacts more epileptiformally active were selected. For convenient, since this point we will refer to the IEEG signals just as EEG signals.

B. State-Space Reconstruction

Some researchers have postulated that the brain is governed by a chaotic dynamics and, thus, have applied nonlinear dynamic tools to analyze EEG data [11], [19]. However, other researchers discard such claims saying that the brain is too complex to be reduced to a system constituted by a handful of state variables. Though arguably the brain dynamics might not be simple [5], in this work, our interests are to look at the EEG data from a chaos theory perspective allowing us to take advantages of nonlinear dynamic tools to analyze the dynamics to measure changes using an evolutionary algorithm (discussed further), just as we can model a signal as a random process and derive statistic measures to describe and characterize it.

Chaos theory [22] states that within a chaotic system, that is, a system displaying apparently disordered random-like data, an underlying order exists. Because of this, one may conjecture that even if precise long-term prediction is impossible, prediction in the short term and with some error allowance may be possible in many systems. Such a property allows us to reconstruct the state-space trajectory of an attractor of the system, which in this case would be a model of the brain that generates the EEG signals.

We can reconstruct the deterministic component of the state trajectory of the EEG signals by taking previous samples of the observable output and creating an artificial state vector with n_e elements (i.e., the number of delays), which we denote λ_k

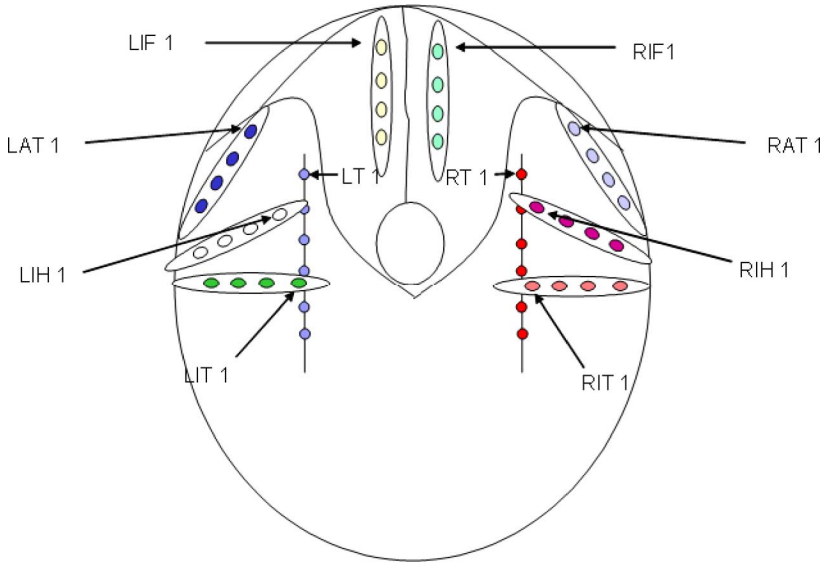


Fig. 2. Diagram of the electrode contact arrangement. The letters on the diagram denote the following: L = left, R = right, T = temporal, I = inferior, A = anterior, H = hippocampal, and F = frontal. For example, LAT stands for left anterior temporal lobe.

(input to the program in Fig. 1), the embedding vector [23]. This process creates a diffeomorphism [9], a smoothly distorted copy of the original trajectory that preserves the dynamics and topological properties of the trajectory, such as number of holes, sharp corners and, thus, dimensional and dynamical properties such as fractal dimension and Lyapunov exponents. In other words, if we calculate a dynamic property (e.g., fractal dimension) of the original trajectory and of its diffeomorphism, the results will be approximates. To have a clear idea, one of the well-known chaotic systems is the Lorenz attractor [22]. The Lorenz attractor's equations, although simple, they produce chaotic behavior. When three-dimension system ($x(t)$, $y(t)$, and $z(t)$) is plotted in a three-dimension space, the trajectory of the system can be observed. What we stated above is that a diffeomorphism of the Lorenz attractor can be constructed using the previous values of an observable measure from system under study. In this case, it is known that the order of the Lorenz's system is 3. Therefore, using a pseudo-state vector formed by a state variable and two delay samples of it, we reconstruct the trajectory. Selecting $x(t)$ as the measure signal, we can reconstruct the smoothly trajectory by setting $x_e(t) = x(t)$, $y_e(t) = x(t-\tau)$, and $z_e(t) = x(t-2\tau)$, where τ is any delay greater than zero (but in practice is best determined from a decorrelation time). In general, n_e, the embedding dimension, needs to be twice the topological dimension to fit the (possibly holey and curvy) manifold, or twice plus one (value known as Takens' bound) that to generically guarantee an embedding.

C. Genetic Programming

Genetic programming algorithm is a paradigm formalized by Koza [13] in which he wanted to probe whether an algorithm using a bunch of mathematical operators could find the solver or set of instructions to solve a given problem. Inspired from genetic algorithms (GA) [13], genetic programming (GP) provides computer programs that are constituted with nodes from the function and terminal set. The terminal set contains the input variables

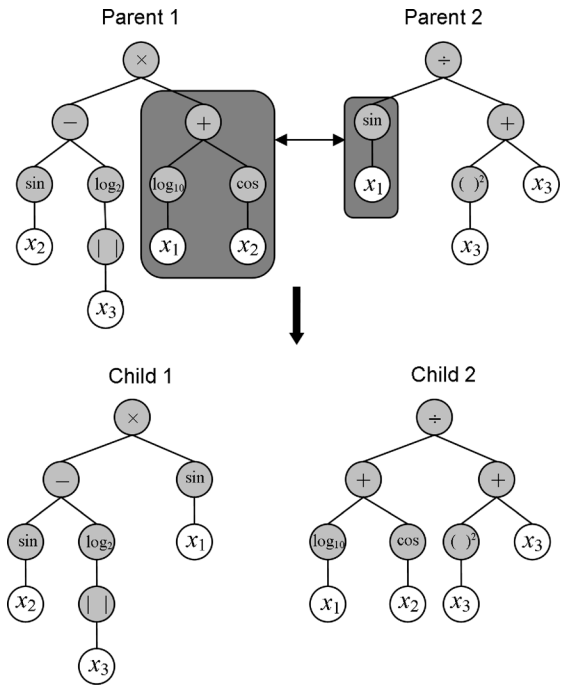


Fig. 3. Crossover: Two node-braches randomly selected in the parent programs are swapped (crossover) to create two new programs (children), which will be evaluated to calculate the fitness.

whereas the function set contains the mathematical operators, which have some number of arguments that they operate on. A typical way to represent the GP-programs is using trees as shown in Figs. 3 and 4. For clarity, let us take as example the tree from Fig. 3 denoted parent 1 (top left). This tree encodes the function: $(\sin(x_2) - \log_2(|x_3|)) \times (\log_{10}(x_1) + \cos(x_2))$, where $+$, $-$, cos, sin, $|\,|$, $\log_2$, and $\log_{10}$ are operators constituting the function set and x_1, x_2, and x_3 constitute the terminal set.

GP uses the genetic operators: selection, crossover, mutation, and reproduction to evolve the new populations attempting to

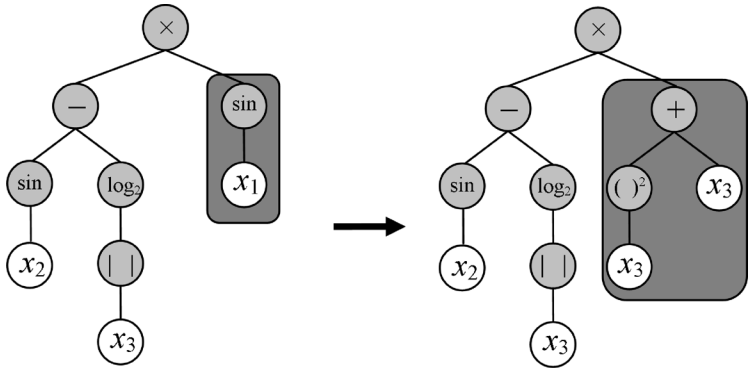


Fig. 4. Mutation: A node-branch is randomly selected in the first program (upper) and substituted by a new branch randomly generated (below) creating a new program or function.

TABLE I
LIST OF FUNCTIONS AND THEIR RESPECTIVE SYMBOLS FOR THE GP ALGORITHM

Function	Arity	Symbol
addition	2	+
subtraction	2	−
multiplication	2	×
division	2	÷
square	1	$(\ )^2$
square root	1	$\sqrt{\ }$
natural logarithm	1	ln
logarithm base 2	1	$\log_2$
logarithm base 10	1	$\log_{10}$
absolute value	1	$\lvert\ \rvert$
sine	1	sin
cosine	1	cos

find the best solver (i.e., function or program). Fig. 3 shows an example where two programs (parents) being crossed over and resulting in two new programs (children). Fig. 4 depicts a program (function) being mutated. The node-branches swapped or mutated were selected randomly, though more sophisticated strategies have been proposed in the literature. These children functions might be potential features to separate our baseline and ictal classes. Whether they are good features or not, it is determined by evaluating an objective function, a function that measures the intended goal (separation of the baseline and ictal classes, for instance) and scores each of these possible solutions. Those that obtain a good score go to the next generation to evolve into new possible solutions. How many and which of these solutions go to the next generation it is dictated by the selection operator. Later, to create the next potential solutions from those functions that came from the previous generation, the crossover and mutation operators are used and so on.

Table I shows the function set or building blocks to be used by the GP algorithm to construct the programs (artificial features). To satisfy the closure property (i.e., the GP algorithm should be able to handle any function with no problem or violation), the square root, logarithm, and division operators are implemented in a protected way. Protected-division works identically to ordinary division except that it outputs one when its denominator input is zero. In order to avoid complex values, the protected square root operator applies an absolute value operator to the input value before the taking the square root. Also, logarithms are protected versions, which output zero when for an argument of zero and apply an absolute value operator to the argument to handle negative arguments.

The terminal set available to the GP algorithm in our application is $\{x[n], x[n-\tau], x[n-2\tau], x[n-3\tau], x[n-4\tau], x[n-5\tau]\}$. From the chaos experimentalist's point of view, the number of reconstructed-dynamic variables in the terminal set is $n_e = 6$, the embedding dimension. Such parameter was set based on prior tests on the limits of these methods by numerically recovering the fractal dimension of known systems from their time-series trajectories at the embedding ranging from 1 to 20 and sufficiently long records (constant: 0; logistic map: 0.5; sinusoid: 1; Hénon map: 1.27; sum of sines on a 2-torus: 2; Lorenz attractor: 2.07; sum of sines on 3-, 4-, and 5-tori: 3, 4, and 5; white noise: monotonic increase). Results are reasonable for objects of dimension up to 3, which places the embedding dimension at 6 or 7, as stated by the Takens' bound.

The delay time τ is determined by looking for the first zero-crossing point of the autocorrelation function of the signal [17] —the time τ after t when $x(t)$ stops "looking like itself." This parameter differs for each patient.

The GP algorithm was extended to allow it to evaluate not just one tree per individual, but multiple trees per individual (i.e., a forest), allowing us to generate a prespecified number of features for a patient simultaneously. All the trees of the individual were evaluated at the same time; however, there was no crossover among trees (features) of the same individual—crossover was done only among homologous trees.

D. Genetic Programming Artificial Features

Having discussed the state-space reconstruction (using delay embedding) and genetic programming algorithm, we proceed to explain our approach to design epileptic seizure detectors using artificial features. Given the state-space trajectory reconstruction (preprocessing stage in Fig. 1) using delays of EEG signal, our approach is to input these pseudo-state vectors to the genetic programming (GP) algorithm (feature creation stage in

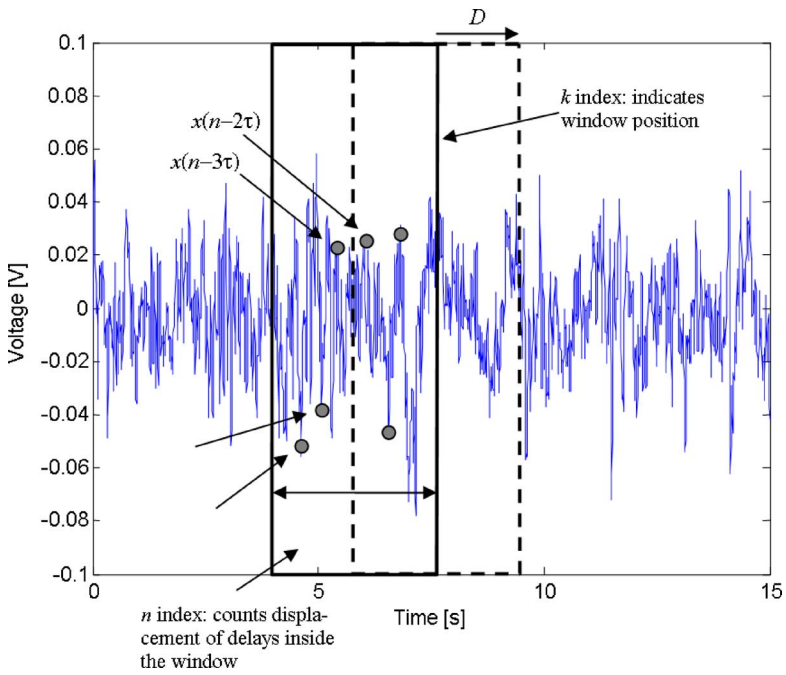


Fig. 5. Sliding window: The sliding window slides (solid rectangle) over an EEG signal. L is the length of the window and D is displacement. The sliding window position is controlled by k, thus, the dashed rectangle, which indicates the next position where the window will be displaced, is the window position $k+1$. The shifting of the delays (indicated by gray dots) inside sliding window is controlled by the index n.

Fig. 1), and by means of the algorithm, to find a transformation, probably nonlinear, that achieves the maximum separability between baseline and ictal data. In other words, the GP algorithm creates a program (artificial feature) that combines the inputs (space-state variables) in a nonlinear way and outputs a function that is intended to separate the baseline and ictal signals such that the performance of the classifier is optimal in the sense of a minimum-error-risk objective function. Mathematically speaking

$$y[k] = \phi(\lambda) \tag{1}$$

where ϕ is a transformation function (program to calculate artificial feature y) designed by the GP algorithm, and

$$\lambda = \{x[n], x[n-\tau], x[n-2\tau], \ldots, x[n-(n_e-1)\tau]\} \tag{2}$$

is the set of delayed samples (the terminal set for the GP) that the GP will use in any combination to construct the artificial features. Therefore, the artificial feature is a function defined ϕ_i: $\mathbf{R}^{ne} \rightarrow \mathbf{R}$.

In evaluating the genetic programming artificial features (GPAF) algorithm, we process the data as viewed through a sliding observation window. Fig. 5 shows an illustration, which presents a sliding window being slid through an EEG segment. The sliding window is defined by two parameters: length of the window, denoted L, and displacement, denoted D. The length L defines the number of points that will be evaluated at the same time. On the other hand, D is defined as the number of new points that will be used in the next evaluation or $D = L - O$, where O is the overlap with the previous window. To reduce variability in the artificial feature signal, we used a summation in each window position, operation that is proportional to the average and, thus, posses the same smoothing property. Recalling (1) and placing the summation operator, the resultant equation is stated as

$$y_i[k] = \sum_{n=1+(k-1)D}^{D(k-1)+L} \phi_i(x[n], x[n-\tau], \ldots, x[n-(n_e-1)\tau]), \quad 1 \le k \le \mathrm{K}. \tag{3}$$

The $y_i[k]$ is the GP artificial feature value of the EEG signal $x(t)$, termed here the *artificial feature time-series*, the subscript i denotes the index of the feature for possible multiple artificial features, n is the original EEG time index that controls the displacement of the delays inside the sliding window, and k is the discrete time unit of the artificial feature time-series (or the index that indicates the next sliding-window position) and goes from 1 to K. The sliding window observes all the L points in the window-position k, whereas the index n is delimited by $1+(k-1)D \le n \le D(k-1)+L$ (superscript and subscript on the summation symbol). Note if n advances by steps of one, and D is chosen to be $\le L$, then all the EEG data available will be used, scrutinizing the data in maximum detail. Fig. 5 illustrates the n and k indexes as well as an instance of delays, with $n_e = 6$, depicted as gray dots inside the window. In other words, (3) is proportional to the average of the GPAF-processed points observed through the sliding window, where the argument of the summation $\phi(\cdot)$ (i.e., the GP artificial feature) is found by the GP algorithm. At each window position k, a number of points, L, will be reduced, by means of (3), to a single point at the output. Thus, if $x(t)$ contains P_{Tot} points, the number of points that

contains the artificial feature time-series is processed by the GP artificial feature is

$$\mathrm{K} = 1 + \left\lfloor \frac{P - L - 1}{D} \right\rfloor. \tag{4}$$

E. Objective Function

As previously mentioned, the GP algorithm relies on an objective (fitness) function to drive the search for best set of artificial features. The ideal seizure detector would be this that achieves zero false negatives and zero false positives, that is, zero errors. Coming from this point, we used the error risk, stated in (5), as the objective function to be optimized [7], [8], an equation derived from the probability error equation, but because the prior probability that a given time window belongs to the seizure class is very small, we can compensate with the term related to the false negative with a large risk factor in order the GP algorithm can pay attention to the number of false negative as well. Otherwise the GP algorithm can find artificial features that classify the baseline data perfectly while obtaining a detector with a high rate of false negatives, making it impractical

$$R_{\mathrm{E}} = P_{\mathrm{FN}} P(\mathrm{S}_T) r_{\mathrm{FN}} + P_{\mathrm{FP}} P(\mathrm{NS}_T) r_{\mathrm{FP}} \tag{5}$$

where P_{FN} is the probability of false negatives (i.e., when the algorithm misses a seizure), P_{FP} is the probability of false positives, and $P(S_T)$ and $P(\mathrm{NS}_T)$ are the prior probabilities of the respective classes (preseizure and baseline). Additionally, r_{FN} is a risk factor associated with missing seizures, and r_{FP} is a risk factor associated with declaring false positives. From experience, we select $r_{\mathrm{FN}} = 0.75$ and $r_{\mathrm{FP}} = 0.25$, yielding

$$R_{\mathrm{E}} = (0.75)\, P_{\mathrm{FN}} + (0.25)\, P_{\mathrm{FP}}. \tag{6}$$

However, clinical trials of seizure detectors/stimulators have shown that assigning risk weights as anywhere from 50%-50% to 75%-25% provide satisfactory results because the effect of the false positives has turned out to be much milder than previously thought.

F. Classifier

As a decision structure for the GPAF algorithm, any classifier may be used. For this paper, we selected the general-purpose k-nearest neighbor (k-NN) classifier [3] (classification stage in Fig. 1). This classifier is nonparametric, nonlinear, and capable of producing multiple thresholds or complicated decision boundaries, making it suitable for n-dimensional, multi-modal problems. In addition, the training process is relatively easy, simply involving giving to the classifier all of the training data.

After the classification, a decision integration window that is slid through the output of the classifier to make an integer decision on when to stimulate the patient was used. If an implantable device (with GPAF features integrated) classified each t s over an incoming EEG signal, the device could not be allowed to stimulate the patient based on the decision that the classifier makes each t s. A longer, fixed-length window is needed, so that it can observe the past evolution of the last t s classification during a defined period to decide whether or not a patient is suffering a seizure. We denote the length of this window as L_{DIW}. This parameter controls a tradeoff between the number of false positives and the detection delays of seizures. Thus, it also controls the number of seizures that are detected; however, because most seizures are detected sooner or later given the sudden and large changes on the EEG signal's amplitude, what is also of great interest is how early can those seizures be detected. This parameter can be fixed as a constant for all patients, that is, the same L_{DIW} value for all the patients. However, this parameter gives flexibility to our approach; therefore, to make the system more "optimal" it is better to tune the parameter for each patient. The L_{DIW} was selected by trial-and-error for each patient.

G. Description of the Seizure Detection Experiments

Since the relative amplitude variability of the EEG signals at the moment of a seizure is high, feature points were generated every 0.125 s. Therefore, the length of the sliding window was set to $D = 50$ (0.25 s) and $L = 200$ (1 s). The training data consisted on three baseline and three ictal epochs for each patient of 2-min duration each (i.e., epoch contains 24 000 points), whereas the validation data is a long EEG archive lasting many hours for each patient. For the case of ictal epochs, the first minute of data before the unequivocal electrograph onset (UEO), which is the point in time marked by a board-certified electroencephalographer with forward and backward paging allowed (non-causal and certain declaration of seizure), was labeled as baseline data and the rest of the epoch (i.e., the minute after the UEO) was labeled as ictal class. Fig. 6 depicts the labeling scheme for an ictal epoch. From the picture, it can be noted that the 1-min EEG data before the UEO has an amplitude similar to that of the baseline data, which amplitude varies from -0.4 mV and 0.4 mV. Given that for the detection case a large part of the information used to detect seizures is provided by the amplitude of the signal, it is reasonable that labeling the whole epoch as belonging to one class (i.e., ictal) would be confusing to the GPAF (or any other) algorithm.

The number of nearest neighbors for the k-NN classifier was set to $k = 5$ (same value used for the benchmark), a value typically used in the literature. For the experiments, we sought to achieve a false negative rate of 10% or less and no more than two false positives per day (or a false positive rate of 0.0833 per hour).

As mentioned previously, the function set for the GP algorithm is listed in Table I, whereas the terminal set is $\{x[n], x[n-\tau], x[n-2\tau], x[n-3\tau], x[n-4\tau], x[n-5\tau]\}$ since n_e, the embedding dimension, was set to six. We set the number of trees for each individual in the GP algorithm to two; that is, the GPAF algorithm created two artificial features for each patient. The maximum depth for a tree was set to 10 and the population size to 1000 individuals. The initial population was generated using the ramped half-and-half method. A crossover operator with a rate of 0.9, fitness proportional selection, and a breeding operator with a rate of 0.1 were used.

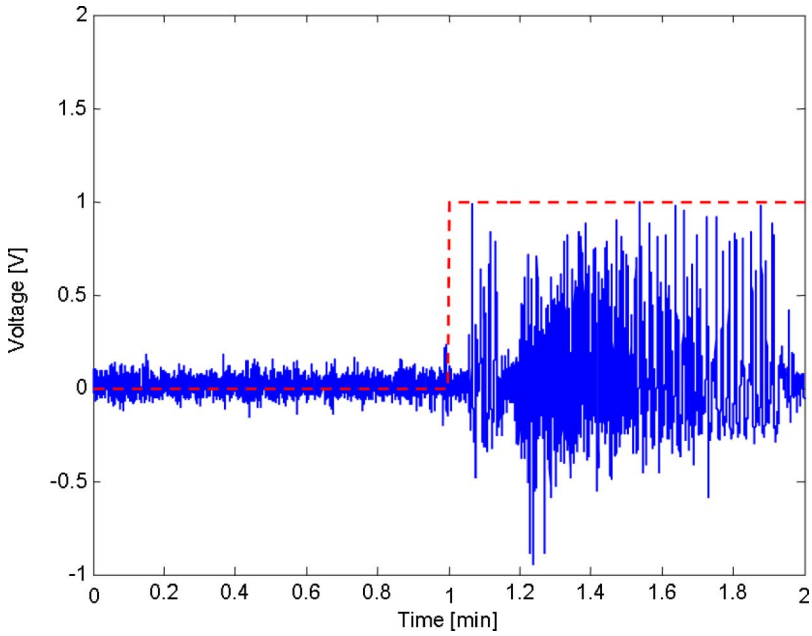


Fig. 6. Classification labeling for a 2-min ictal epoch used for training purposes. The first minute of data was labeled as baseline (dashed line is set at zero) and the other minute of data was labeled as ictal data (dashed line is set to one). The UEO occurs exactly at the minute.

H. Illustrated Example

Fig. 7 shows all the components comprising the GPAF algorithm and the steps involved in the evaluation of one individual. In the figure, we provide an example where two artificial features are being sought (thus, an individual corresponds to a two-tree "forest"). The first step is the generation of the initial population of artificial features (**I**). The artificial features are transformed from vectors to trees for comprehension purposes (**II**). Such trees encode the artificial features. From these, the artificial functions are translated into mathematical expressions (the arguments of the summation operators) to be evaluated (**III**). Later, the EEG signal, segmented with the sliding window (**IV**) is processed by the artificial features generating the artificial signals (**V**). Later, these artificial data are input to the k-NN classifier to execute the categorization task (**VI**). In the illustration, label "0" denotes the baseline samples and "1" denotes the ictal samples. In the training phase (i.e., when artificial features are being designed), we could evaluate either the training data or a checking data set (a data set different from the training data intended to resemble the validation data) to calculate the classification error, the performance measure that will drive the GP search. It should be noted that this entire illustration is just to evaluate one individual (i.e., one forest). After all individuals in the initial population are evaluated, the genetic operators take over to create the next population of artificial features. The errors are calculated again. This process is repeated until either a performance goal is reached or the maximum number of generation is met.

I. Benchmark Detector

Results from the GPAF algorithm are compared with the results provided by a handcrafted feature proposed by Esteller *et al.* [5]. In their work, they used the line length feature, which proved to be one of the best man-crafted features for detection of epileptic seizures. This feature achieved a high accuracy of detection and a low false alarm rate. In this work, we implemented the line length feature as a benchmark detector

$$LL[k] = \sum_{n=1+50(k-1)}^{50k+150} |x[n] - x[n-1]|. \tag{7}$$

After the EEG signals were processed by the above conventional feature, the k-NN classifier was used to perform the categorization task (i.e., detection). Parameter values k and L_{DIW} were the same used for the GPAF features for each patient.

III. Results

A. Patient-Specific Detectors

1) Patient A: This patient suffered six simplex and complex partial seizures found coming from the left anterior hippocampus through 45.0 hr of EEG recording. The adjacent channels more epileptoformally active were LT2-LT3 (i.e., left temporal lobe 2-left temporal lobe 3). Based on autocorrelation analysis of the signal, the delay was set to $\tau = 19$ (or 0.095 s) and terminal set was $\{x[n], x[n-19], x[n-38], x[n-57], x[n-76], x[n-95]\}$.Equation (8) shows the artificial features found by the GPAF algorithm for the detector of this patient's seizures. Table II provides the results for the validation data using the GPAF and line length features, where FP denotes the number of false positives, FPH is the rate of false positive per hour, FN denotes the number of false negatives, L_{DIW} is the length of the decision integration window, ADT denotes the average delay of detection time, Sz is the number of seizures that occurred throughout the EEG archive, and HP is the number of hours of EEG recordings processed for validation. This patient had a

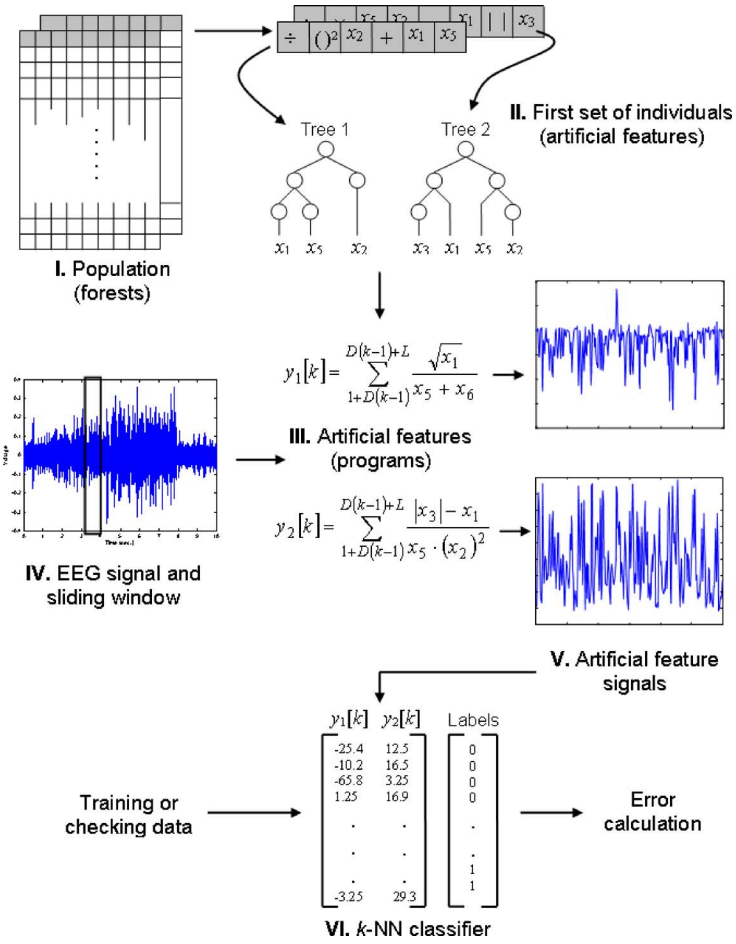


Fig. 7. Detailed illustration of the GPAF algorithm evaluating one individual (i.e., a forest).

TABLE II
RESULTS FOR VALIDATION DATA FOR PATIENT A

Feature	FP	FPH	FN	L_{DIW}	ADT	Sz	HP
GPAF	6	0.133	0	3.75 s	15.67 s	6	45.0
Line length	5	0.111	0	3.75 s	7.67 s	6	45.0

high average of detection time for the GPAF equations because one seizure was detected with a delay of 35.75 s

$$y_1[k] = \sum_{n=1+50(k-1)}^{50k+150} \frac{(x[n-76])^2}{\log(x[n-19])}$$
$$y_2[k] = \sum_{n=1+50(k-1)}^{50k+150} x[n-95] \cdot x[n-76]. \quad (8)$$

2) Patient B: Through 46.2 hr of EEG recording, this patient suffered five complex partial seizures, and its most active pair was RIT3-RIT4. The delay was $\tau = 11$ (or 0.055 s) and terminal set was $\{x[n], x[n-11], x[n-22], x[n-33], x[n-44], x[n-55]\}$. Equation (9) shows the equations designed by the GPAF algorithm. Additionally, Table III provides the results obtained from the validation of the GPAF equations. The table shows that just one false positive occurred. Also, it was able to detect all five seizures that occurred. In addition, the artificial detector was able to achieve a low average detection delay

$$y_1[k] = \sum_{n=1+50(k-1)}^{50k+150} \log_{10}\left(x[n-11] - x[n-22] + \sqrt{x[n-11]}\right)$$
$$y_2[k] = \sum_{n=1+50(k-1)}^{50k+150} \log_2\left(\frac{\log_2(x[n-11])}{x[n-33] + x[n-44]}\right). \quad (9)$$

For this patient, a good average of detection delay was achieved because this patient had a characteristic pattern called "beta-buzz" in which the amplitude increases rapidly and then drops suddenly before the UEO. For this patient, the GPAF features never under-performed the line length detector, and outperformed it on average time to detection (as noted from Table III).

3) Patient C: This patient had an archive of 65.7 hr of EEG recording in which eleven complex partial, multifocal clinical

TABLE III
RESULTS FOR VALIDATION DATA FOR PATIENT B

Feature	FP	FPH	FN	L_{DIW}	ADT	Sz	HP
GPAF	1	0.022	0	3.75 s	1.15	5	46.2
Line length	1	0.022	0	3.75 s	4.35 s	5	46.2

TABLE IV
RESULTS FOR VALIDATION DATA FOR PATIENT C

Feature	FP	FPH	FN	L_{DIW}	ADT	Sz	HP
GPAF	0	0	0	3 s	1.93 s	11	65.7
Line length	0	0	0	3 s	–1.82 s	11	65.7

TABLE V
RESULTS FOR VALIDATION DATA FOR PATIENT D

Feature	FP	FPH	FN	L_{DIW}	ADT	Sz	HP
GPAF	0	0	0	2.75 s	2.1 s	11	156.4
Line length	0	0	0	2.75 s	4 s	11	156.4

seizures were diagnosed coming from right and left inferior neocortex, left temporal neocortex, and left hippocampus. The most active channel pairs were LT2-LT3 and RIF1-RIF2. The delay was $\tau = 19$ (or 0.095 s) and terminal set was $\{x[n], x[n-19], x[n-38], x[n-57], x[n-76], x[n-95]\}$. The pair of artificial features found by the GPAF algorithm is in (10). Table IV shows the results for the validation data, where zero false positives and, thus, a zero false positive rate per hour occurred and all eleven seizures were detected. Finally, though not better than the line length, the artificial equations produced a low average of detection delay

$$y_1[k] = \sum_{n=1+50(k-1)}^{50k+150} \left(x[n] - \frac{\sqrt{x[n-76]}}{x[n-95]} \right)^2$$
$$y_2[k] = \sum_{n=1+50(k-1)}^{50k+150} \left| \frac{(x[n-19] - x[n])^2}{|\log_{10}(\log_{10}(x[n-38]))|} \right|. \quad (10)$$

4) Patient D: This patient has 156.4 hr of EEG recording in which eleven clinical complex partial seizures were found coming from the right inferior frontal region. To derive the bipolar EEG data, channels RIF2-RIF3 were selected. The delay was $\tau = 9$ (or 0.045 s) and terminal set was $\{x[n], x[n-9], x[n-18], x[n-27], x[n-36], x[n-45]\}$. Equation (11) shows the artificial functions found by the GPAF algorithm. As seen in Table V, the GPAF equations and the line length both performed well, producing no false positives or false negatives, and the average detection delay (ADT) with GPAF was lower than that with line length

$$y_1[k] = \sum_{n=1+50(k-1)}^{50k+150} |x[n-9] - x[n]|$$
$$y_2[k] = \sum_{n=1+50(k-1)}^{50k+150} |x[n]|. \quad (11)$$

5) Patient E: Patient E has the longer EEG archive, lasting 197.3 hr, in which 15 complex partial seizures coming from the left hippocampus and left anterior temporal cortex were found, and its selected adjacent pair channel was LT2-LT3. The delay was set $\tau = 9$ (or 0.045 s) and, thus, its terminal set was $\{x[n], x[n-9], x[n-18], x[n-27], x[n-36], x[n-45]\}$. The pair of artificial features is presented in (12). Table VI provides the results for the artificial features. The algorithm produced five false positives, providing a false positive rate of 0.025 per hour, which is a very good rate. All seizures were detected with the GPAF features producing a slightly better average delay of detection although this difference does not appear to be statistically significant.

$$y_1[k] = \sum_{n=1+50(k-1)}^{50k+150} x[n-18] \cdot x[n-36]$$
$$y_2[k] = \sum_{n=1+50(k-1)}^{50k+150} x[n] + x[n-9] + x[n-36] \cdot x[n-45]. \quad (12)$$

6) Patient F: The entire EEG recording of this patient lasts 66.2 hr, during which seven complex partial clinical seizures were identified coming from bilateral regions. LT2-LT3 and RIT2-RIT3 were selected as the most active pair. The delay was $\tau = 20$ (or 0.1 s) and terminal set was $\{x[n], x[n-20], x[n-40], x[n-60], x[n-80], x[n-100]\}$. Equation (13) shows the artificial features designed by the GPAF algorithm. Table VII shows results obtained with the GPAF equations. For this patient, the GPAF-designed equations produced three false positives. The average detection delay was low. On the other hand, the algorithm failed to detect one seizure previously reported. However, the GPAF algorithm could detect three clinical seizures and one subclinical seizure that had not been previously reported; whereas the line length feature detected an extra subclinical seizure (that is, the GPAF

TABLE VI
RESULTS FOR VALIDATION DATA FOR PATIENT E

Feature	FP	FPH	FN	L_{DIW}	ADT	Sz	HP
GPAF	5	0.025	0	4.25 s	10.9 s	15	197.3
Line length	5	0.025	0	4.25 s	11.05 s	15	197.3

TABLE VII
RESULTS FOR VALIDATION DATA FOR PATIENT F

Feature	FP	FPH	FN	L_{DIW}	ADT	Sz	HP
GPAF	3	0.045	2	3.75 s	-9.25 s	12	66.2
Line length	4	0.060	1	3.75 s	5.33 s	12	66.2

TABLE VIII
RESULTS FOR VALIDATION DATA FOR PATIENT G

Feature	FP	FPH	FN	L_{DIW}	ADT	Sz	HP
GPAF	11	0.072	2	2.75 s	5.81 s	32	153.8
Line length	81	0.527	0	2.75 s	3.40 s	32	153.8

features did not detect the additional subclinical seizure). The clinical seizures lasted roughly 1 min, whereas the subclinical seizures lasted 16–18 s

$$y_1[k] = \sum_{n=1+50(k-1)}^{50k+150} \log\left(\sqrt{x[n-100]} - x[n-20]\right)$$
$$y_2[k] = \sum_{n=1+50(k-1)}^{50k+150} \left|\frac{x[n-100]}{x[n]\cdot \log_{10}(x[n-60])}\right|. \quad (13)$$

7) Patient G: This patient had available 153.8 hr of EEG for the training and validation procedures. LIT1-LIT2 were the adjacent channels more active. This patient suffered 31 complex partial seizures coming from the left hippocampus and left inferior temporal neocortex, of which 20 were subclinical. We included the subclinical seizures so that the GPAF algorithm could design a detector that also warns about such seizures because it is known that they are also disruptive to the brain [13]. Subclinical seizures are epileptic seizures, but of duration no more than 30 s and without overt, clinically evident symptoms—i.e., they are detected only with EEGs.

The delay was $\tau = 15$ (or 0.045 s) and hence the terminal set was $\{x[n], x[n-15], x[n-30], x[n-45], x[n-60], x[n-75]\}$. Equation (14) shows the artificial features found. Table VIII provides the results for the validation data. Clearly, this pair of artificial features is achieving good performance. The two seizures missed by the GPAF equations were both subclinical seizures. However, an additional unreported subclinical seizure was detected by both methods. Additionally, a low false positive rate was achieved. In contrast, the line length feature had a high false positive rate. In fact, the number of false positives for the line length feature was larger because it misclassified almost 8 hr of EEG recording. If the parameter L_{DIW} was increased for the line length feature to reduce the number of false positives, several seizures were missed

$$y_1[k] = \sum_{n=1+50(k-1)}^{50k+150} x[n-75]\cdot x[n-60] + \frac{x[n-60]}{x[n-30]}$$
$$y_2[k] = \sum_{n=1+50(k-1)}^{50k+150} \frac{(x[n-60])^2}{\ln(x[n-45])}. \quad (14)$$

B. Generic Detector

A generic detector, such as the line length feature, is desired because it can be applied on any patient without the need of extracting data to design a detector for the seizures of the patient at hand. By just giving to the detector a small amount of baseline and ictal data, the detector should able to detect most seizures and at the same time to have a low false positive rate. In this subsection, we show a general detector designed by the GPAF algorithm.

For this experiment, we only used one baseline epoch and one ictal epoch of 2-min duration from each available patient. Therefore, the training data set consisted of seven baseline and seven ictal epochs (a total of 28 min of data). As an ictal epoch for patient G, we selected a subclinical seizure epoch, using the clinical seizures of the other patients in the training database to emulate a clinical seizure for patient G. Otherwise we would not have a subclinical epoch as example in the training data.

In searching for a generic model (detector), we also fixed the delay parameter by obtaining an average delay from the delays previously calculated. The average delay was 16. Therefore, the terminal set for the GP algorithm was $\{x[n], x[n-16], x[n-32], x[n-48], x[n-64], x[n-80]\}$. The LDIW parameter was adjusted for each patient and was not necessarily the same value used for the customized detector in the previous section for the patient-specific cases. The parameters for the GP algorithm and

TABLE IX
RESULTS FOR VALIDATION DATA FOR THE GENERIC DETECTOR

Patient	FP	FPH	FN	L_{DIW}	ADT	Sz	HP
A	3	0.067	0	3.75 s	15.08 s	6	45.0
B	2	0.043	0	3.75 s	0.1 s	5	46.2
C	0	0	0	3.25 s	0.16 s	11	65.7
D	0	0	0	2.5 s	1.55 s	11	156.4
E	25	0.159	0	6.25 s	19.13 s	15	197.3
F	8	0.121	1	6.25 s	16.83 s	13	66.2
G	2	0.013	2	2.75 s	6.22 s	32	153.8
Overall	5.71	0.058	0.429	4.07 s	8.44 s	93	730.6

k-NN classifier were set as before. Equation (15) shows the artificial equations found by the GPAF algorithm.

Table IX provides the results for the seven patients using the validation data, their respective whole EEG archives. The results were acceptable for most patients, having similar performance to the customized features in the previous section. Five patients out of seven met the specifications of a 0.0833 of false positive rate and less than 10% of false negative, as shown in (15) at the bottom of the page.

For patient E, a 0% false negative rate was achieved; however, this patient had an "outbreak" of EEG prodromes in the form of false starts and delta trains, which caused all of the false positives that it had. Because of this, the length of the decision window was increased to filter out several false positives; however, this decision affected the average detection delay.

Patient F had a rate of 9.09% false negatives, counting three clinical seizures that were detected, and which had not been reported before. On the other hand, this patient had many epileptiform discharges (spike-and-wave patterns) that lasted only a few seconds. These electrical discharges made the general detector fire, producing several false positives. In addition to the three clinical and two subclinical unreported seizures previously found, the detector found an additional subclinical seizure, which was not detected by the customized detector.

Finally, a similar situation occurred with the results for patient G. Two subclinical seizures were missed, yielding a false negative rate of 6.45%. However, a previously unreported subclinical seizure was found. A low false positive rate per hour was achieved. Additionally, a series of 19 epileptiform discharges—spike-and-wave EEG patterns—were also detected. These triggered the detector.

IV. DISCUSSION

The GPAF algorithm can produce or design sets of equations with similar or better performance to that of a conventionally human-crafted feature—in this case, the line length. For the patient-specific artificial features, just four seizures were missed (with three being subclinical seizures), for a rate of 4.35% false negatives (88 of 92 seizures caught, counting the previously unreported ones), which more than satisfies the requirement set, being no more than 10%. Also, in 730.6 hr (roughly 30.4 days) of EEG recordings, just 23 false positives were signaled, for a rate of 0.0315 false positives per hour. The false positive requirement was 0.0833 false positive per hour so the system exceeded the expectation. The line length feature achieved an average 0.1314 false positives per hour. However, this value was affected mostly by the results in patient G. Also, it missed just one previously reported seizure on patient F, yielding a false negative rate of 1.09%. Note, however, that GPAF was using two equations (artificial features) while the line length is just one equation, allowing the GPAF equations more flexibility.

A generic detector for all patients was also successful. This model was sufficiently general to allow use of a fixed (average) delay parameter. The overall false positive rate per hour was a low 0.055, and only three seizures were missed, of which two were subclinicals. However, this model was also able to detect an additional subclinical seizure in patient F. This yielded a low false negative rate of 3.22% (90 seizures of 93 were detected, after counting the seizures that had not been previously reported). This general detector was created by the GPAF algorithm with just one baseline and one ictal epoch of 2–min duration from each patient, a very limited dataset (due to computational expense) compared with the entire body of data available. Time (generations) limits on the evolution of GPAF features were also imposed, for the same reason.

The number of false positives and false negatives (and, thus, the average of detection delay) is affected by the length of the decision window. Although the length can be fixed, it seemed beneficial to fix the length for each patient. Additionally, the design of artificial features from the GPAF algorithm is affected by

$$y_1[k] = \sum_{n=1+50(k-1)}^{50k+150} \left(x[n-64] - x[n-80]\right)^2 + \frac{x[n] \cdot \sin\left(\sin\left(x[n]\right)\right) \cdot x[n-48]}{\left(x[n-16]\right)^2 - \log_{10}\left(x[n-16]\right)} - \log_2\left(x[n-64]\right)$$

$$y_2[k] = \sum_{n=1+50(k-1)}^{50k+150} x[n] + x[n-80] + x[n-32] \cdot x[n-16] - \cos\left(x[n-48]\right) \tag{15}$$

where we start to label the ictal class. From the outset, we were labeling the ictal epochs 1 min before the UEO in order to give to the GPAF algorithm a way to track some sort of signal that indicates that an epileptic seizure was imminent. However, artificial features then produced too many false positives, because part of the data was in conflict with the post-UEO data, as was explained in Section II-G. Therefore, we decided to start labeling data as ictal class only after UEO. After that, good artificial features were obtained. However, we also learned that the classification should not start for all patients at UEO. For example, patient B has a characteristic pattern before UEO, known as beta-buzz and which occurs 1 s or so before the UEO, in which amplitude collapses and frequency increases considerably. This pattern was precisely the one that made the detector fire, allowing an early detection of the seizures. However, in general, the length of the decision integration window, the classification labeling, and the possible feature complexity provided by the GP algorithm provide much flexibility to design good artificial features. First, the GP algorithm tries to explore the search space looking for the "optimal" artificial features, based on the data and the classification labeling. Later, the points misclassified can be corrected using the decision integration window, using its length parameter to control the tradeoff between false positives and negatives. This architecture allows us to design detectors that pay attention to the EEG patterns and specifics of each patient.

Finally, this work has shown that the GPAF algorithm is not only able to create artificial features with no known physical meaning, but also is able to find some features that are known to be good features—specifically, this was the case for patient D delay set to nine (instead of one as in the line length feature). Although not the central object of this study, the results for patient D confirmed the claims by Esteller *et al.* [5] in the sense that the line length feature is a good and efficient feature for detection of epileptic seizure onsets.

V. Conclusion

We presented the GPAF algorithm, an algorithm intended to overcome most of the drawbacks of conventional strategies for feature gathering, along with the space-state reconstruction preprocessing to detect epileptic seizure onsets, demonstrating its use in several experiments. These experiments yielded detectors that had low false positive rates per hour and low false negative rates (in total, 88 of 92 seizures were detected). In addition to customized equations that were compared with the line length feature and had similar or better results than that benchmark, we also presented a generic detector that showed good performance by detecting 90 seizures (clinical and subclinical, including the seizures discovered during the experiments) out of 93, and by being able to detect an extra seizure that was not detected by either the other GPAF-generated detectors or by the benchmark.

We cannot claim that the GPAF algorithm outperformed the line length feature; rather, the GPAF algorithm did better for some patients and the line length feature did better for other patients. However, the GPAF algorithm has a desirable property: it provides room for improvement. By providing more data, adding a term in the fitness function that considers in some better way the detection delay, and using high-end computers to enable use of more data and more thorough search, there is a high probability that the GPAF algorithm will find better artificial features that respond to the characteristics of a particular patient than does one universal feature. Additionally, we can increase the number of artificial features, for additional discriminating power.

We may expect to extract more information and, thus, construct a better detector by allowing the GPAF algorithm to select the channel or set of channels maximizing the fitness function of the algorithm instead of selecting the most active bipolar channel as we did. A similar approach was employed by D'Alessandro in [2]. A well-known problem with providing more information is that it increases the dimensionality of the data and traditional methods are unable to handle such high-dimensional space. The GPAF algorithm "degrades gracefully" under such circumstances since one of its components is a genetic programming algorithm, a global and stochastic algorithm, which with enough resources (i.e., time and computational power), can find a reliable and practical solution, that scale well with dimensionality the dimension of the data.

References

[1] H. Chen, S. Zhong, and D. Yao, "Detection singularity value of character wave in epileptic EEG by wavelet," in *Proc. IEEE Int. Conf. Communications, Circuits and Systems and West Sino Expositions 2002*, Jun.–Jul. 29–1, 2002, vol. 2, pp. 1094–1097.

[2] M. D'Alessandro, "The utility of intracranial EEG feature and channel synergy for evaluating the spatial and temporal behavior of seizure precursors," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, 2001.

[3] R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. New York, NY: Wiley, 2001.

[4] M. D'Alessandro, "Wavelet neural networks for EEG modeling and classification," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, 1995.

[5] R. Esteller, J. Echauz, T. Tcheng, B. Litt, and B. Pless, "Line length: An efficient feature for seizure onset detection," in *Proc. 23rd Annu. EMBS Int. Conf.*, 2001, pp. 1707–1709.

[6] R. Esteller, G. Vachtsevanos, J. Echauz, M. D'Alessandro, C. Bowen, R. Shor, and Litt, "Fractal dimension detects seizures onset in mesial temporal lobe epilepsy," in *Proc. 1st Joint BMES/EMBS Conf.: Serving Humanity, Advancing Technology*, Atlanta, GA, Oct. 13–16, 1999, p. 442.

[7] H. Firpi, "Genetically found, neurally computed artificial features with applications to epileptic seizure detection and prediction," masters thesis, Univ. Puerto Rico-Mayagüez, Mayagüez, 2001.

[8] ——, "On prediction and detection of epileptic seizures by means of genetic programming artificial features," Ph.D. dissertation, Michigan State University, Lansing, 2005.

[9] R. Gilmore and M. Lefranc, *The Topology of Chaos: Alice in the Stretch and Squeezeland*. New York: Wiley., 2002.

[10] R. L. Haupt and S. E. Haupt, *Practical Genetic Algorithms*. New York: Wiley, 1998.

[11] B. H. Jansen, "Quantitative analysis of electroencephalograms: Is there chaos in the future?," *Int. J. Biomed. Computing*, vol. 27, pp. 95–123, 1991.

[12] J. R. Koza, *Genetic Programming: On Programming of Computers by Means of Natural Selection*. Cambridge, MA: MIT Press, 1992.

[13] S. A. Lee, D. D. Spencer, and S. S. Spencer, "Intracranial EEG seizure-onset patterns in neocortical epilepsy," *Epilepsia*, vol. 41, no. 3, pp. 297–307, 2000.

[14] M. M. Mirbagheri, K. Badie, R. M. H. Golpayegani, and M. A. Ahmadi, "A neural network approach to EEG classification for the propose of differential diagnosis between epilepsy and normal EEG states," in *Proc. Annu. Int. Conf. IEEE Engineering in Medicine and Biology Society*, 1992, vol. 14, pp. 2649–2650.

[15] C. Niederhoefer, F. Gollas, A. Chernihovskyi, K. Lehnertz, and R. Tetzlaff, "Detection of seizure precursors in the EEG with cellular neural networks," *Epilepsia*, vol. 45, no. Suppl. 7 (abstract), p. 245, 2004.

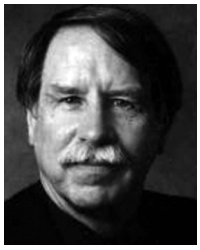
[16] D. Olsen, R. Lesser, J. Harris, R. Webber, and J. Cristion, "Automatic Detection of Seizures Using Electroencephalographic Signals," US Patent 5 311 876, 1994.

[17] S. J. Orfanidis, *Optimum Signal Processing: An Introduction*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[18] I. Osorio, M. G. Frei, and S. B. Wilkinson, "Real time automated detection and quantitative analysis of seizures and short term predictions of clinical onset," *Epilepsia*, pp. 615–627, 1998.
[19] W. S. Pritchard and D. W. Duke, "Measuring chaos in the brain: A tutorial review of nonlinear dynamical EEG analysis," *Int. J. Neuroscience*, Apr. 1992, to be published.
[20] H. Qu and J. Gotman, "A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: possible use as a warning device," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 2, pp. 115–122, Feb. 1997.
[21] L. Szilágyi, Z. Benyó, and S. M. Szilágyi, "A new method for epileptic waveform recognition using wavelets decomposition and artificial neural networks," in *Proc. 2nd Joint EMBS/BMES Conf.*, Houston, TX, Oct. 23–26, 2002, pp. 2025–2026.
[22] J. C. Sprott, *Chaos and Time-Series Analysis*. New York: Oxford Univ. Press, 2003.
[23] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical Systems and Turbulence*. Berlin, Germany: Springer-Verlag, 1981, vol. 898, Warwick 1980 Lecture Notes in Mathematics, pp. 336–381.
[24] R. Tetzlaff, C. Niederhofer, and P. Fischer, "Feature extraction in epilepsy using a cellular neural network based device – First results," in *Proc. Int. Symp. Circuit and Systems 2003*, May 25–28, 2003, vol. 3, pp. III-850–III-853.

Hiram Firpi (S'99–M'06) earned the B.S.E.E. degree from the Polytechnic University of Puerto Rico, San Juan, Puerto Rico, in 1999, a M.S.E.E. degree from University of Puerto Rico, Mayaguez, Puerto Rico, in 2001, and the Ph.D. degree in electrical engineering from Michigan State University, East Lansing, in 2005.

In 2005, he became a Postdoctoral Fellow with the Intelligent Control System Laboratory at the Georgia Institute of Technology, Atlanta. He is now a Postdoctoral Fellow with the Indiana University-Perdue University at Indianapolis, where he will be developing and implementing intelligent methods for bioinformatic and computational biology problems. His research interests range from pattern recognition and machine learning tools to biomedical and control applications.

Dr. Firpi is a member of the Tau Beta Pi honor society.



Erik D. Goodman received the B.S. degree in mathematics and the M.S. degree in systems science, from Michigan State University (MSU), East Lansing, in 1966 and 1968, respectively. He received the Ph.D. degree in computer and communication sciences in 1971 from the University of Michigan, Ann Arbor, in 1971.

He became an Assistant Professor of Electrical Engineering and Systems Science in 1972, Associate Professor in 1978, and Professor in 1984, all at MSU, where he was also appointed Professor in Mechanical Engineering in 1992. He directed the Case Center for Computer-Aided Engineering and Manufacturing from 1983 to 2002, and MSU's Manufacturing Research Consortium from 1993 to 2003. He has co-directed MSU's Genetic Algorithms Research and Applications Group (GARAGe) since its founding in 1993. He has been doing research in evolutionary computation since 1970, and has also published extensively in ecosystem modeling and simulation and in manufacturing process simulation and verification. His current research involves automated design of dynamic systems using genetic programming and of mechanical structures using hierarchical and hybrid genetic algorithms.

Dr. Goodman is Chair of the Executive Board of the International Society for Genetic and Evolutionary Computation, a Senior Member of the Society of Manufacturing Engineers, and the chair of ACM SIGEVO, the Special Interest Group on Genetic and Evolutionary Computation.

Javier Echauz (S'93–M'95) received the B.S.E.E. degree in 1988 from University of Puerto Rico-Mayagüez (UPRM), Mayagüez, Puerto Rico, and M.S.E.E. and Ph.D. degrees in the areas of computational intelligence, control systems, nonlinear dynamics, and quantitative EEG analysis from the Georgia Institute of Technology, Atlanta, in 1989 and 1995, respectively.

Before cofounding Atlanta-based startups IntelliMedix, Inc. and BioQuantix Corp., he was Associate Professor and Associate Director of the Electrical and Computer Engineering Department at UPRM. He was also Senior Research Scientist at NeuroPace, Inc., contributing to the development of an implantable brain neurostimulator. He has coauthored some 40 technical papers, 2 book chapters, and 6 patents. His research interests since 1997 have centered around intelligent devices for epilepsy.