Classification of EEG Signals Using Fractional Calculus and Wavelet Support Vector Machine

Aaruni V. C. *, Harsha A. † and Liza Annie Joseph
Rajagiri School of Engineering & Technology, Kakkanad
Cochin, Kerala- 682039.
Email: *aarunivc@gmail.com, †harsha.appukutan@gmail.com

***Abstract*—The behavior of many physical and biological processes and systems can be described satisfactorily by fractional order models. A new method, termed fractional linear prediction (FLP) based on fractional calculus, is used to model ictal and seizure-free EEG signals. Through numerical simulations it is demonstrated that, the EEG signal can be modeled accurately, by using a few integrals of fractional orders as basis functions. The parameters obtained from modeling are used for analysis and classification using support vector machines (SVM). It is found that improvements in classification accuracy is possible by using wavelet support vector machines using wavelet kernel functions such as Mexican hat wavelet and Morlet wavelet.**

***Index Terms*—Electroencephalogram (EEG) signal, Fractional linear prediction (FLP), Epileptic seizure classification, support vector machine (SVM), SVM kernel, wavelet SVM (WSVM).**

I. Introduction

Neural activity in the human brain starts from the early stages of prenatal development and are electrical in nature. Both the brain function and the status of the whole body are represented by these signals. Electroencephalography is one of the most widely used methods to assess brain activity. The simplicity, low cost and high temporal resolution of EEG makes it a more popular tool than the other methods like magnetoencephalogram (MEG) and functional magnetic resonance imaging (fMRI). The proper understanding of the neuronal functions and biosignals are of vital importance for the detection, diagnosis and treatment of brain disorders [1].

Epilepsy is a group of chronic neurological disorders. Its main feature is the recurrence of seizures that result in irregular disturbance of brain function and causes the appearance of characteristics rhythms (spikes, polyspikes, spike-and-wave complex). Due to the presence of these spikes, ictal EEG signals are more impulsive in nature than the seizure-free EEG signals. The detection of epileptic seizures in the EEG signals is an important part in the diagnosis of epilepsy. Epileptic seizure detection by visual scanning of EEG signal may not be accurate and requires much time. So, for efficient diagnosis of epilepsy, successful EEG modeling is required [2].

For automatic detection of epileptic seizures, several parameters extracted from EEG signals can act as very useful diagnostic features. The commonly used features for detection and classification of epileptic seizure EEG signals are the spectral parameters based on the Fourier transform. However, the analysis based on Fourier transform assumes that the analyzed signal is stationary. Various studies have shown that the EEG signal is non-stationary in nature [3], [4]. There are several time-frequency domain based methods like short time Fourier transform, the multifractal analysis, empirical mode decomposition [5], the wavelet transform and the multi-wavelet transform for epileptic seizure detection from EEG signals. Autoregressive models are also used to find the seizure locations of epileptic EEG signals. Epileptic seizure can be detected successfully using linear prediction error energy also. The method used in this paper is inspired from the fact that linear predictive coding (LPC) [6] can be used for EEG signal modeling.

As we know, epileptic EEG signals have sharp transitions and LPC based on ordinary integer order differential equations cannot trace these small changes efficiently. Hence, in this paper, instead of integer order models, we use fractional order models. This gives a better EEG signal modeling. The technique used is fractional linear prediction (FLP) which is based on fractional calculus. The new approach requires a lesser number of model parameters and signal analysis is much better. By combining wavelet techniques with SVMs we get wavelet SVM that gives better classification accuracy.

The rest of the paper is organized as follows: Section II deals with the database and materials used in this paper. The method followed in this paper is described in Section III. In Section IV, we have the experimental results and discussions for modeling and classification of EEG signals. Finally Section V concludes the paper.

II. DATASET

An EEG dataset, which is publicly available online in [7] is used in this work. In this dataset, there are five subsets denoted as Z, O, N, F, and S. Each subset contains 100 single-channel EEG signals of which each has 23.6 s duration. The sampling frequency of EEG signals in the dataset is 173.61 Hz. The subsets Z and O have been recorded extracranially, whereas the subsets N, F, and S have been recorded intracranially. The subset S contains seizure activity, selected from all recording sites exhibiting ictal activity. In this work, the seizure-free class is formed by combining the subsets Z, O, N and F. The ictal class consists of the subset S.

III. METHODOLOGY

A flowchart of the proposed method is shown in Fig. 1. We pass the EEG signal through an FLP filter. Through a least-squares approach, the FLP coefficients of the signal are calculated. The signal is modeled using these coefficients. In this work, modeling is done on the first 800 samples of each signal. There occurs some difference between the actual signal and the modeled signal which is termed as prediction error. After signal modeling, we estimate the prediction error energy. The signal energy is also calculated and we repeat this procedure for the entire dataset. The sharp changes that appear in ictal EEG signals can not be modeled accurately due to the low pass nature of FLP and hence the prediction error increases. Hence they require higher order FLP for modeling, whereas seizure-free EEG signals require lower order. Modeling error will be higher for ictal EEG signals, if we use the same order FLP filter for modeling both ictal and seizure-free EEG signals.

For classifying EEG signals, a classification system can be developed based on the modeling error energy and signal energy. The prediction error energy and signal energy of 50% of the data from each category are used as features to train a SVM. The trained SVM is used for classification. For classification of EEG signals, the remaining 50% of feature data are used. Next, we combine the wavelet technique with SVM to get wavelet support vector machines (WSVM). Classification accuracy is tested using wavelet kernel functions such as Mexican hat and Morlet.

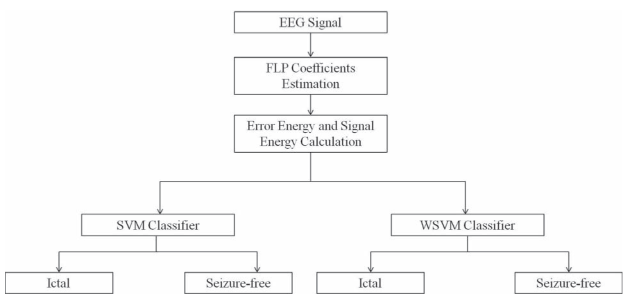

Fig. 1. Flowchart

A. *Fractional Linear Prediction (FLP)*

The fractional calculus is a name for the theory of integrals and derivatives of arbitrary order which unify and generalize the notions of integer-order differentiation and n-fold integration [8], [9]. In the recent years, fractional calculus has been successfully applied to many areas such as automatic control and signal processing. It has been found that speech signals can be modeled more effectively using fractional order modeling techniques than linear prediction (LP) techniques [10]. Here, FLP is used to give a more accurate representation of EEG signal [11]. Fractional derivatives can be defined in many ways. Riemann-Liouville [12] is the most commonly used definition. According to this, the fractional derivative of order o of a function $x(t)$ can be expressed as follows:

$$\frac{d^o x(t)}{dt^o} = \frac{1}{\Gamma(m-o)} \frac{d^m}{dt^m} \int_0^t \frac{x(\tau)}{(t-\tau)^{o-m+1}} d\tau \qquad (1)$$

where $m-1$ is an integer, $m-1 < o \leq m$, and the Euler's Gamma function $\Gamma(z)$ is defined as:

$$\Gamma(z) = \int_0^\infty e^{-x} x^{z-1} dx \qquad (2)$$

The Grunwald-Letnikov approximation of the fractional derivative is commonly used [13] for performing numerical simulations on a computer. The fractional derivative of order o of a function $x(t)$ can be expressed by Grunwald-Letnikov approximation as :

$$D^o x(t) = \lim_{h \to 0} h^{-o} \sum_{p=0}^{\lceil t/h \rceil} (-1)^p \binom{o}{p} x(t-ph) \qquad (3)$$

In traditional continuous-time linear system, we represent output signal as a linear combination of integer order derivatives of input signal. Similarly, in FLP, we can express a predicted EEG signal as a linear combination of its fractional derivatives as shown:

$$\hat{x}(n) = \sum_{k=1}^{Q} \lambda_k D^{o_k} x(n) \qquad (4)$$

Note that here a negative value of o_k corresponds to fractional integral of order o_k. For noise immunity and numerical stability, we may recast the above equation using Eq. (3) as follows:

$$\hat{x}(n) = \sum_{k=1}^{Q} \lambda_k D^{o_k} x(n) = \sum_{k=1}^{Q} \alpha_k \Psi_k(n) \qquad (5)$$

where α_k are the required FLP parameters. These parameters can be determined by minimizing the prediction error energy which is given by:

$$\epsilon = \sum_{n=0}^{N-1} (e(n))^2 \qquad (6)$$

where $e(n)$ is the prediction error:

$$e(n) = x(n) - \hat{x}(n) \qquad (7)$$

and N is the number of samples in the signal. We need to determine α_k while minimizing the prediction error energy ϵ. For convenience, we rewrite the above equations in vector-matrix notation. The sequence corresponding to fractional integral $\Psi_k(n)$ is denoted by $N \times 1$ column vectors ρ_k and the required coefficients are denoted by the column vector g. Through a least square approach, we then need to solve the following equation to get α_k :

$$g = \left(\Lambda^T \Lambda\right)^{-1} \Lambda^T x \qquad (8)$$

where

$$\Lambda = [\rho_1 \ \rho_2 \ldots \rho_Q] \qquad (9)$$

Thus, we get the required FLP coefficients and using these we can model our signal.

TABLE I
CONTINGENCY TABLE

Predicted Class	Actual Class	
	Ictal	Seizure-free
Ictal	TP	FP
Seizure-free	FN	TN

B. Classification using SVM

Here, SVM is used for EEG signal classification since it is the simplest method for classification. Support vector machines [14] are machine learning method based on statistical learning theory. Strong robustness to noise and generalization to unseen data are expected from SVM and this subject is the most important drawback of artificial neural networks. One can easily understand the basic principle of an SVM for a two-dimensional case. Suppose that we need to classsify a series of data points into two different classes [15]. The SVM method provides a boundary between the two classes such that the distance between the boundary and the nearest data point in each class is maximal. We call the nearest data points as support vectors.

A hyperplane needs to be created to do linear separation in higher dimensions, if the separation into two classes is not possible by a linear boundary. An SVM uses a device called kernel mapping to map the data in input space to a high-dimensional feature space in which the problem becomes linearly separable. This is achieved by using a transformation $T(x)$ that maps the data from input space to feature space. If a kernel function

$$M(x,y) = T(x).T(y) \tag{10}$$

is used to perform the transformation, then the basic form of SVM can be expressed as follows:

$$g(x) = sign\left(\sum_{i=1}^{l} \beta_i y_i M(x, x_i) + a\right) \tag{11}$$

We vary the kernel functions and their parameter values used for training the SVM to get the highest accuracy. Here, the kernel functions used are linear, polynomial and RBF kernel functions. For polynomial kernel, the parameter that can be changed is the polynomial order. Scaling factor (sigma) is the parameter that can be varied in the radial basis function kernel. The performance of the method is evaluated through SVM classification plots and classification test performance determined by accuracy (Acc), sensitivity (SEN) and specificity (SPE) values based on True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN) for the set of classified data. They are defined as:

$$SEN = \frac{TP}{TP + FN} \times 100 \tag{12}$$

$$SPE = \frac{TN}{TN + FP} \times 100 \tag{13}$$

$$ACC = \frac{Currectly\,Classified}{Total} \times 100 \tag{14}$$

The classification performance can be represented using a contingency table as shown in Table I.

C. Classification using WSVM

Since the wavelet technique shows promise for both nonstationary signal approximation and classification it is valuable for us to study the problem of whether a better performance could be obtained if we combine the wavelet technique with SVMs [16]. Wavelet decomposition is a powerful tool for approximation. The wavelets are commonly used basis functions for time-frequency analysis of nonstationary signals. In wavelet analysis, a signal or function is approximated by a family of functions generated by dilations and translations of a function called the mother wavelet $\phi(x)$ [17]. The dilated and translated mother wavelet function is a rapidly decreasing oscillation function given by

$$\phi_{a,c}(x) = \frac{1}{\sqrt{a}}\phi\left(\frac{x-c}{a}\right) \tag{15}$$

where $a, c \in \Re, a > 0, a$ is the dilaiton parameter, and the analysis wavelet function is centered at time c. The multidimensional wavelet function can be defined as the product of one-dimensional wavelet functions as:

$$\phi_d(x) = \prod_{k=1}^{d} \phi(x^k) \tag{16}$$

where $\phi(x)$ is the single dimension wavelet function.

Here, the wavelet kernel has the same expression as a multidimensional wavelet function. The goal of the WSVMs is to find the optimal approximation or classification in the space spanned by multidimensional wavelets or wavelet kernels. We use function handles to define new kernel functions in MATLAB. If $x, x_i \in \Re^d$, then the dot product wavelet kernels are obtained as [4], [5], [16], [18]

$$K(x, x_i) = \prod_{k=1}^{d} \phi\left(\frac{x^k - c^k}{a}\right)\left(\frac{x_i^k - c_i^k}{a}\right) \tag{17}$$

and translation-invariant wavelet kernels are expressed as:

$$K(x, x_i) = \prod_{k=1}^{d} \phi\left(\frac{x^k - x_i^k}{a}\right) \tag{18}$$

The wavelet kernel based SVM decision function is obtained

$$f(x) = sign[\sum_{i=1}^{N} \alpha_i y_i \prod_{k=1}^{d} \phi(\frac{x^k - x_i^k}{a}) + b] \tag{19}$$

where x_i^k is the kth component of ith training data. We use the WSVM using the Morlet and Mexican hat wavelet kernel for classification of EEG signals into ictal and seizure-free categories. The kernel function that can be defined from the Mexican hat mother wavelet,
$\phi(x) = (1 - x^2)e^{-\frac{x^2}{2}}$ as

$$K(x, x_i) = \prod_{k=1}^{d} \left[1 - \frac{(x^k - x_i^k)^2}{a^2}\right] e^{-\frac{\left\|x^k - x_i^k\right\|^2}{2a^2}} \tag{20}$$

Similarly, the kernel function that can be defined for the Morlet mother wavelet,
$\phi(x) = cos[w_0 x]e^{-\frac{x^2}{2}}$ as

$$K(x, x_i) = \prod_{k=1}^{d} cos\left[w_0 \frac{(x^k - x_i^k)}{a}\right] e^{-\frac{\left\|x^k - x_i^k\right\|^2}{2a^2}} \qquad (21)$$

The optimal kernel parameters can be selected by trial and error method. The classification test performance is evaluated for WSVMs also.

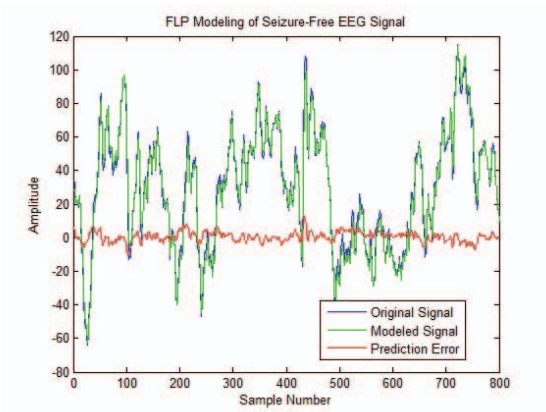

Fig. 2. FLP Modeling of Seizure-free EEG Signal

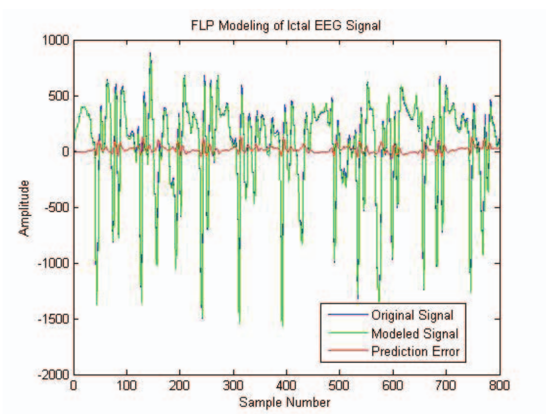

Fig. 3. FLP Modeling of Ictal EEG Signals

IV. RESULTS

A. Modeling

EEG signal modeling is done using FLP. To compare the performance of FLP with the integer order linear prediction, modeling the EEG signals using linear predictive coding is also done. A 12^{th} order linear prediction filter is used here. The modeling results obtained using FLP with only two coefficients based on the fractional integrals 0.15 and 0.3 are shown in Fig. 2 & Fig. 3.

The prediction errors for both linear prediction and fractional linear prediction were computed. In Fig. 4, the prediction errors for FLP and LPC modeling of seizure-free EEG data are shown. The Fig. 5 shows the prediction errors for ictal EEG data modeling. It can be observed from the figures that in FLP model, we get smoother error signal compared to LPC model.

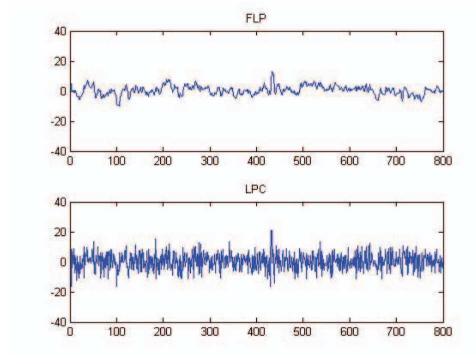

Fig. 4. Prediction Errors for FLP and LPC Modeling of Seizure-free EEG

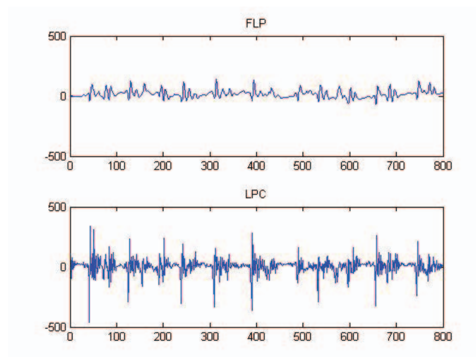

Fig. 5. Prediction Errors for FLP and LPC Modeling of Ictal EEG

B. Classification

SVM classification was done using linear, polynomial and RBF kernels. Classification is also done by using wavelet support vector machines (WSVMs). The wavelet kernel functions used are Morlet and Mexican hat wavelet kernels. Classification accuracy is improved by using WSVMs. The contingency table for Morlet kernel is shown in Table II.

TABLE II
CONTINGENCY TABLE FOR MORLET KERNEL

Predicted Class	Actual Class	
	Ictal	Seizure-free
Ictal	46	6
Seizure-free	4	194

The classification of data into ictal and seizure-free classes for Morlet wavelet kernel is shown in Fig. 6. It is clear from the figure that the proposed method can be used as a diagnostic tool for detecting ictal EEG signals. The classification accuracies for different kernel functions for each data set are summarized in Table III.

The classification test performance of the SVM classifier for different kernel functions are given in Table IV. The maximum

TABLE III
CLASSIFICATION ACCURACY FOR DIFFERENT KERNEL FUNCTIONS AND EEG DATASETS

Kernel Function	Accuracy (%)				
	F	N	Z	O	S
Linear	92	96	100	100	86
Polynomial (order 3)	86	96	100	98	94
RBF($\sigma = 0.3$)	86	96	100	98	94
Mexican Hat(a=1.1)	92	98	100	100	86
Morlet (a=.3,$\omega_0 = 0.75$)	90	98	100	100	92

TABLE IV
CLASSIFICATION TEST PERFORMANCE FOR DIFFERENT KERNEL FUNCTIONS

Kernel Function	Sensitivity	Specificity	Accuracy
Linear	86%	97%	94.8%
Polynomial(Order 3)	94%	95%	94.8%
RBF ($\sigma = 0.3$)	94%	95%	94.8%
Mexican Hat(a=1.1)	86%	97.5%	95.2%
Morlet(a=.3,$\omega_0 = 0.75$)	92%	97%	96%

classification accuracy is given by the WSVM using Morlet wavelet as the kernel function which gives 96% accuracy with parameters $a = .3$ and $\omega_0 = 0.75$. Through numerical simulations it is found that the maximum classification accuracy for LPC modeling is attained for RBF kernel which gives a maximum of only 94.4% .

V. CONCLUSIONS

Fractional linear prediction is a novel approach that utilizes fractional integrals as basis functions to model the EEG signal more effectively. It is found that FLP based modeling outperforms the commonly used LPC-based modeling techniques. One major application of this modeling is the classification of EEG signals into ictal and seizure-free categories. The classification of EEG data using prediction error energy and signal energy as parameters to the SVM has proved to be successful. Further improvements in classification accuracy is obtained by using wavelet support vector machines. Wavelet SVMs with Morlet wavelet kernel gives the maximum classification accuracy of 96%. Moreover, the accurate EEG modeling can help in detection of epilepsy, sleep disorders and other neurodegenerative diseases. FLP modeling can be used for EEG compression also. Hence, FLP shows considerable promise to become an important tool for biomedical signal processing applications.

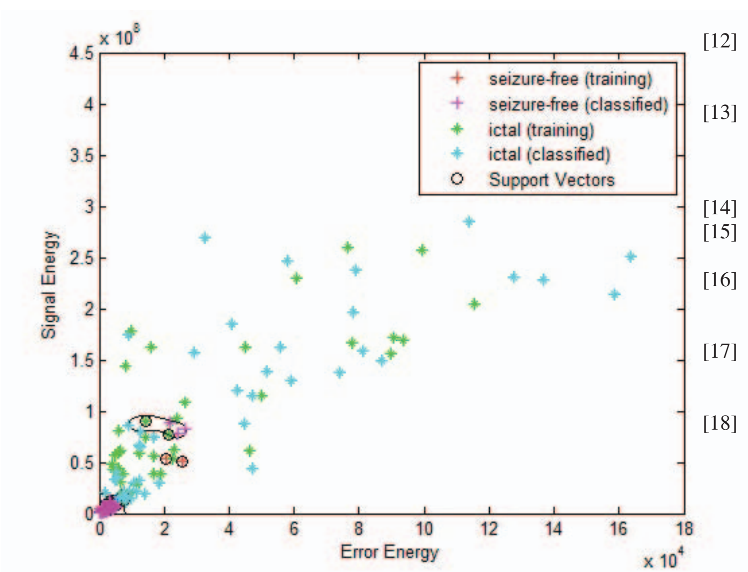

Fig. 6. SVM Classification Plot Using Morlet Wavelet Kernel

REFERENCES

[1] C. Guerrero-Mosquera, A. M. Trigueros, and A. Navia-Vazquez, *Epilepsy - Histological, Electroencephalographic and Psychological Aspects*. InTech, 2012, ch. EEG Signal Processing for Epilepsy.

[2] L. D. Iasemidis., D.-S. Shiau, W. Chaovalitwongse, J. C. Sackellares, P. M. Pardalos, J. C. Principe, P. R. Carney, A. Prasad, B. Veeramani, and K. Tsakalis, "Adaptive epileptic seizure prediction system," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 5, pp. 616–627, 2003.

[3] R. B. Pachori and P. Sircar, "Eeg signal analysis using fb expansion and second-order linear tvar process," *Signal Processing*, vol. 88, no. 2, pp. 415–420, 2008.

[4] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Automatic seizure detection based on time-frequency analysis and artificial neural networks," *Computational Intelligence and Neuroscience*, vol. 2007, pp. 1–13, 2007.

[5] V. Bajaj and R. B. Pachori., "Classification of seizure and nonseizure eeg signals using empirical mode decomposition," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 6, pp. 1135–1142, 2012.

[6] T. Lajnef, S. Chaibi, A. Kachouri, and M. Samet, "Epileptic seizure detection using linear prediction filter," in *11th International conference on Sciences and Techniques of Automatic control computer engineering*, 2010, pp. 1–9.

[7] www.meb.unibonn.de/epileptologie/science/physik/eegdata.html.

[8] H. Sheng, Y. Chen, and T. Qiu, *Fractional Processes and Fractional-Order Signal Processing Techniques and Applications*. London-Springer Verlag, 2012.

[9] lgor Podlubny, *Fractional Differential Equations*. Academic Press, 1999, vol. 198.

[10] K. Assaleh and W. M. Ahmad, "Modeling of speech signals using fractional calculus." IEEE, 2007, pp. 1–4.

[11] V. Joshi, R. B. Pachori, and A. Vijesh, "Classification of ictal and seizure-free eeg signals using fractional linear prediction," *Biomedical Signal Processing and Control*, vol. 9, pp. 1–5, 2013.

[12] R. Magin, M. D. Ortigueira, I. Podlubny, and J. Trujillo, "On the fractional signals and systems," *Signal Processing*, vol. 91, no. 3, p. 350–371, 2011.

[13] M. D. Ortigueira and J. J. Trujillo, "Generalized gl fractional derivative and its laplace and fourier transform," in *Proceedings of the ASME 2009 International Design Engineering Technical Conferences Computers and Information in Engineering Conference*, 2009, pp. 1–5.

[14] http://svms.org/tutorials/.

[15] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Disc.*, vol. 2, no. 2, pp. 1–47, 1998.

[16] L. Zhang, W. Zhou, and L. Jiao, "Wavelet support vector machine," *IEEE Trans.Syst.Man,and Cybern.B,Cybern*, vol. 34, no. 1, pp. 34–39, 2004.

[17] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans.Inform.Theory*, vol. 36, pp. 961–1005, 1990.

[18] M. Zavar, S. Rahati, M. R. Akbarzadeh-T, and H. Ghasemifard, "Evolutionary model selection in a wavelet-based support vector machine for automated seizure detection," *Expert Syst. Appl.*, vol. 38, pp. 10 751–10 758, 2011.