Science Journal of Circuits, Systems and Signal Processing
2021; 10(2): 25-37
http://www.sciencepublishinggroup.com/j/cssp
doi: 10.11648/j.cssp.20211002.11
ISSN: 2326-9065 (Print); ISSN: 2326-9073 (Online)

Jacobi Polynomial Transforms-Based Entropy Measures for Focal and Non-Focal EEG Signals Discrimination Using Kernel Machines

Laurent Chanel Djoufack Nkengfack[1, 2, *], Daniel Tchiotsop[2], Romain Atangana[3], Beaudelaire Saha Tchinda[2], Valérie Louis-Door[4], Didier Wolf[4]

[1]Department of Physics, University of Dschang, Dschang, Cameroon

[2]University Institute of Technology Fotso-Victor, University of Dschang, Bandjoun, Cameroon

[3]Higher Teacher Training College (HTTC) of Bertoua, University of Ngaoundéré, Bertoua, Cameroon

[4]ENSEM de Lorraine, University of Lorraine, Nancy, France

Email address:

laurentdjoufack@gmail.com (L. C. D. Nkengfack), daniel.tchiotsop@gmail.com (D. Tchiotsop), azongmeromain@gmail.com (R. Atangana), stblaire@yahoo.fr (B. S. Tchinda), Valerie.Louis@univ-lorraine.fr (V. Louis-Door), Didier.Wolf@univ-lorraine.fr (D. Wolf)

[*]Corresponding author

To cite this article:

Laurent Chanel Djoufack Nkengfack, Daniel Tchiotsop, Romain Atangana, Beaudelaire Saha Tchinda, Valérie Louis-Door, Didier Wolf. Jacobi Polynomial Transforms-Based Entropy Measures for Focal and Non-Focal EEG Signals Discrimination Using Kernel Machines. *Science Journal of Circuits, Systems and Signal Processing.* Vol. 10, No. 2, 2021, pp. 25-37. doi: 10.11648/j.cssp.20211002.11

Received: July 14, 2021; **Accepted**: July 23, 2021; **Published**: August 18, 2021

Abstract: Electroencephalogram (EEG) remains the primary technique in the diagnosis and localization of partial epilepsy seizures. Despite the advent of modern neuroimaging techniques, the use of EEG signals for locating epilepsy-affected brain areas is still convenient. That is why during these last decades, several computer-aided detection (CAD) methodologies have been proposed to detect and discriminate focal (F) EEG signals, and hence locate epileptogenic foci. In this impetus, this paper applied Jacobi polynomial transforms (JPTs)-based entropy measures to analyze the complexity and discriminate the bivariate focal (F) and non-focal (NF) EEG signals. Jacobi polynomial transforms namely discrete Legendre transform (DLT) and discrete Chebyshev transform (DChT) are applied to separate F and NF EEG signals into their different rhythms. Furthermore, entropy measures like approximate entropy (ApEn), sample entropy (SampEn), permutation entropy (PermEn), fuzzy entropy (FuzzyEn) and increment entropy (IncrEn) are extracted. For direct discrimination between F and NF EEG signals, extracted entropies are combined to define different features vectors that are fed as inputs of two kernel machines namely the least-squares support vector machine (LS-SVM) and simple multi-layer perceptron neural network (sMLPNN). Experimental results demonstrated that our methodology achieved the highest performance of 98.33% sensitivity, 98.00% specificity, and 98.17% accuracy in discriminating F and NF EEG signals with sMLPNN classifier. In addition, our methodology will be useful to clinicians in providing an accurate and objective paradigm for locating epilepsy-affected brain areas.

Keywords: Electroencephalogram (EEG) Signals, Jacobi Polynomial Transforms (JPTs), Entropy Measures, Bivariate Focal (F) EEG, Epileptogenic Focus, Kernel Machines

1. Introduction

Electroencephalogram (EEG) signals are recordings of the brain's electrical activity. One of the main diagnostic applications of EEG is in epilepsy [1]. Epilepsy is broadly separated into two categories namely focal (F) or localized and diffuse or generalized epilepsy. F epileptiform discharges affect a partial area of the brain. Around 20% of patients with generalized epilepsy and 60% of patients with localized epilepsy develop resistance to drugs and undergo surgery [2]. Generally, non-treated F epilepsy can progress to generalized epilepsy. Hence, the identification of the F EEG signals can be helpful to locate the epileptogenic focus before pre-surgical evaluation.

Despite the advent of modern neuroimaging techniques

like positron emission tomography (PET), magnetic resonance imaging (MRI), and single-photon emission computed tomography (SPECT), the use of EEG signals for locating epilepsy-affected brain areas is still convenient. Moreover, EEG signals enable higher temporal resolution and hardware costs are significantly lower for EEG sensors. Recently, several computer-aided detection (CAD) methods have been proposed to detect and discriminate F and NF EEG signals, and hence locate epileptogenic focus [3-11]. In addition, it is noted that nonlinear features are generally used for the detection of F EEG signals. Rajeev Sharma et al. [4] defined an integrated index for the identification of F EEG signals using discrete wavelet transform (DWT) and entropy measures. Entropy features are extracted based on DWT separation of the EEG signals and fed into the LS-SVM classifier which achieved the highest average classification accuracy of 84%, sensitivity of 84% and specificity of 84%. Manish Sharma et al. [5] proposed a framework based on orthogonal wavelet filter banks and computed various entropies from the wavelet coefficients of the signals as features given to the LS-SVM for the discrimination of F and non-focal (NF) EEG signals with the highest classification accuracy of 94.25%, and a 91.95% sensitivity and 96.56% specificity. Abhijit et al. proposed a tunable-Q wavelet transform-based multivariate sub-band fuzzy entropy with application to discrimination of F and NF types of EEG signals [6]. Applied to different time-segmented F and NF EEG signals, the Kruskal-Wallis statistical test of their *Q*-based multivariate sub-band fuzzy entropies was more significant for longer-duration EEG signals. Finally, these multivariate fuzzy entropy features were fed to the random forest and least-squares support vector machine (LS-SVM) classifiers, and their method has achieved the highest classification accuracy of 84.67% in discriminating F and NF EEG signals with LS-SVM classifier. Arunkumar et al. [8] used entropies such as approximate entropy (ApEn), sample entropy (SampEn) and Reyni's entropy as features for discrimination of F and NF EEG signals. These entropy features were fed into six different classifiers such as naïve bayes classifier (NBC), radial basis function (RBF), support vector machine (SVM), k-nearest neighbor (KNN) classifier, non-nested generalized exemplars classifier (NNge) and best first decision tree (BFDT) classifier. It was found that NNge classifier gave the highest accuracy of 98%, sensitivity of 100% and specificity of 96%.

However, it is firstly observed that most of the reported F and NF EEG signals discrimination systems have a limited success rate and kernel machines aim to be more accurate than others classifiers machines. In addition, the ability to discern levels of complexity within biological data sets has become increasingly important even if some models extend to be very complex for practical applications on EEG signals classification. On the other hand, more efforts have been focused on the association of EEG rhythms extraction methods and entropy measures to discriminate F and NF EEG signals. Also, some of these discrimination systems do not relate the extracted rhythms to the ones defined in the literature in terms of spectral coefficients as shown by Djoufack et al. [12-13]. In short, despite many techniques used, polynomial transforms are not yet associated with entropy measures for the purpose of F and NF EEG signals discrimination even if it is already shown that the physical interpretation of the spectral coefficients leads to a new issue for automatic diagnosis in epilepsy [12, 13]. In this impetus, this paper aims to develop Jacobi polynomial transforms (JPTs)-based entropy measures like approximate entropy (ApEn), sample entropy (SampEn), permutation entropy (PermEn), fuzzy entropy (FuzzyEn) and increment entropy (IncrEn) for F and NF EEG signals discrimination. JPTs namely discrete Legendre transform (DLT) and discrete Chebychev transform (DChT) are applied to separate F and NF EEG signals into their different rhythms before computing entropy measures. Different feature combinations are formed from these entropies and fed into the LS-SVM for F and NF EEG signals discrimination.

The rest of the paper is organized as follows. Section 2 describes the EEG database used. The EEG rhythms separation using JPTs technique is briefly presented. Algorithms of different entropy measures and kernel machines are also described. Section 3 presents and discusses the results obtained for the discrimination of F and NF EEG signals. Finally, Section 4 summarizes the study with some concluding remarks.

2. Materials and Methodology

This section describes the set of materials and the methodology used to develop our automated discrimination framework, as depicted in Figure 1. Our proposed discrimination framework applies the JPTs to separate EEG rhythms before computing entropy measures that are used to define inputs of kernel machines.

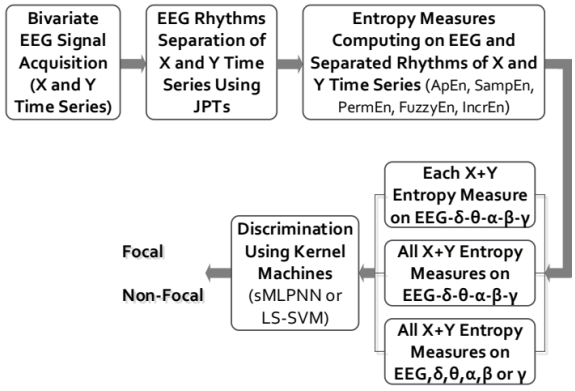

Figure 1. *Block diagram of the proposed F and NF EEG signals discrimination framework.*

2.1. Data Description

The data consists of bivariate F and NF EEG signals in the X and Y time series obtained from the publicly available Bern-Barcelona EEG database [14]. The data are collected from five patients suffering from longstanding pharmaco-resistant temporal lobe epilepsy. Each set (F and NF)

contains 3750 EEG derivations recording with a sampling frequency of 512 Hz. The duration of each derivation is 20 s and leads to 10240 samples. Figure 2 presents an example of EEG signals of the data.

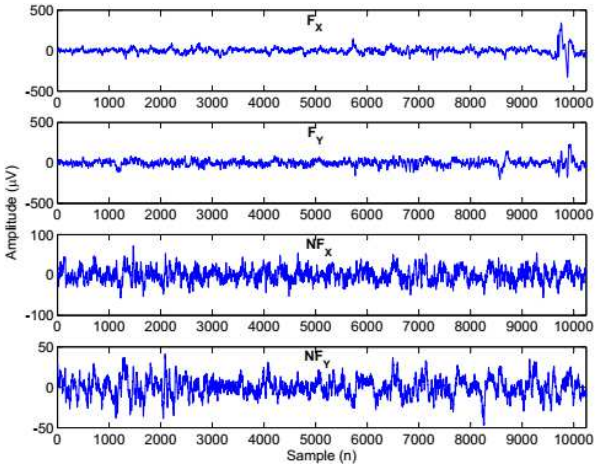

Figure 2. *Exemplary of the Bern-Barcelona EEG signals [14].*

2.2. EEG Rhythms Separation Using JPTs

EEG signals are non-stationaries and typically described in terms of rhythmic activities. Efficient separation of the EEG rhythms can be done based on the projection of the EEG signals into polynomial bases. The proposed scheme is recently briefly described as follows [12]:

Decompose the EEG signal as a set of Jacobi spectral coefficients using DLT and DChT equations in (1) and (2), respectively;

$$\begin{cases} \alpha_k = \dfrac{2k+1}{M(M+1)} \displaystyle\sum_{j=0}^{M} \frac{L_k(x_j)}{\left[L_M(x_j)\right]^2} S(x_j) \;, ? \; k = 0,1,...,M-1 \\ \alpha_M = \dfrac{1}{M+1} \displaystyle\sum_{j=0}^{M} \frac{S(x_j)}{L_M(x_j)} \;, k = M \end{cases} \tag{1}$$

$$\begin{cases} \alpha_0 = \dfrac{1}{M+1} \displaystyle\sum_{j=1}^{M+1} S(x_j) \\ \quad = \dfrac{1}{M+1} \displaystyle\sum_{j=1}^{M+1} S\left\{\cos\left[\frac{(2j-1)}{2(M+1)}\pi\right]\right\} \\ \alpha_k = \dfrac{2}{M+1} \displaystyle\sum_{j=1}^{M+1} S(x_j)\cos\left[\frac{k(2j-1)}{2(M+1)}\pi\right] \\ \quad = \dfrac{2}{M+1} \displaystyle\sum_{j=1}^{M+1} S\left\{\cos\left[\frac{(2j-1)}{2(M+1)}\pi\right]\right\} * \cos\left[\frac{k(2j-1)}{2(M+1)}\pi\right], \quad k = 1,2,3,...,M \end{cases} \tag{2}$$

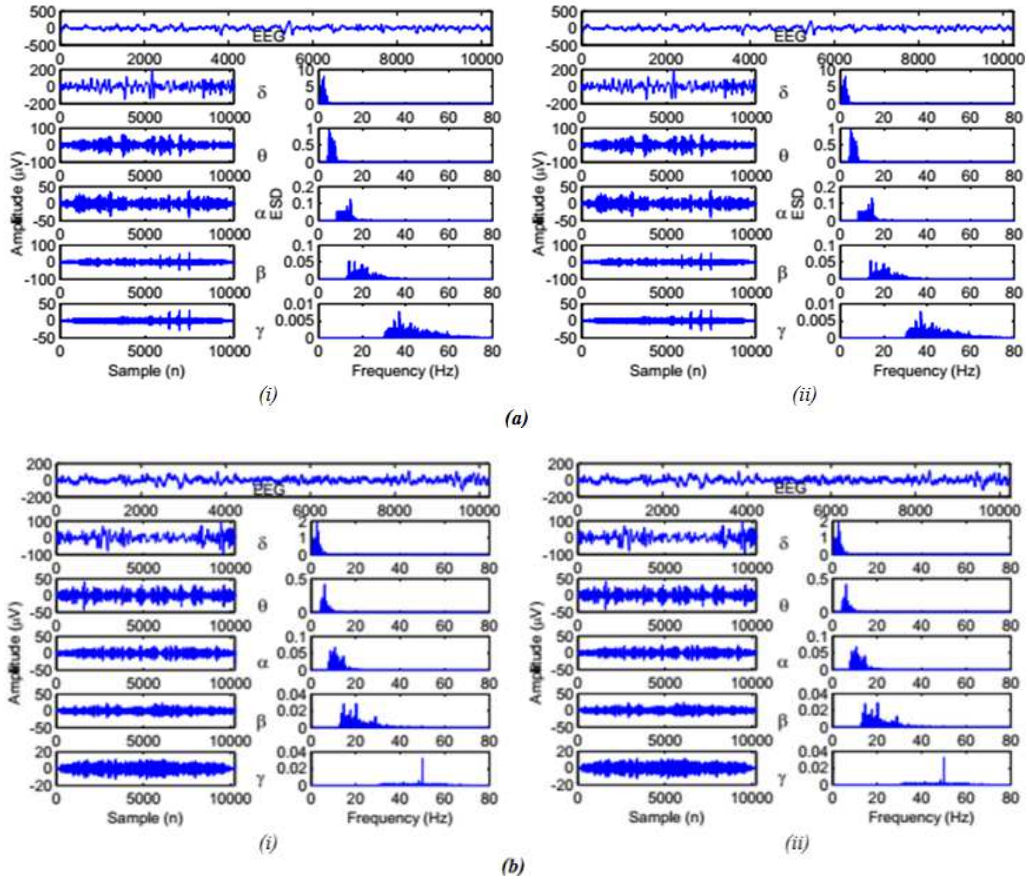

Figure 3. *EEG rhythms separation of (a) a F and (b) a NF X time-series EEG signals using (i) DLT and (ii) DChT. From top to bottom we have the original EEG signal, and the delta, theta, alpha, beta and gamma rhythms, respectively.*

Where the coefficient α_k is the projection of the signal S on the Jacobi base component $J_k^{\alpha,\beta}$, and M is the order of approximation.

Select the spectral coefficients that correspond to each EEG rhythm using the Fourier transform;

Therefore, using the corresponding spectral coefficients for reconstruction with equation (3) provides the rhythms separation of the EEG signal.

$$s(x)=\sum_{k=0}^{M}\alpha_k J_k^{\alpha,\beta}(x) \tag{3}$$

Examples of X time series of F and NF EEG rhythms separation using DLT and DChT are presented in Figure 3.

2.3. Entropy Measures

The notion of entropy is extremely discussed in the literature during this last decade. Characteristics of F and NF EEG signals can be gathered using non-stationary features as entropy measures. After processing the data with the JPTs, EEG signals and their separated rhythms can be represented by typical entropies that are representative of the two classes of bivariate signals. Entropies are mathematical algorithms created to measure the repeatability or predictability of time series. This paper exploits different entropy measures depicted in the following.

2.3.1. Approximate Entropy (ApEn)

Approximate entropy (ApEn) was developed by Steven Pincus to quantify the regularity or predictability of a time series [15-17]. Unlike Shannon's entropy, ApEn accounts the temporal order of points in a time sequence and therefore it is considered as a measure of randomness [17].

Given a time-series data $\{u(i), 1\le i\le N\}$ of length N, the necessary steps involved in the computation of ApEn are provided below.

Step 1: Fix the vector length m as a positive integer, and the relative tolerance limit r as a positive real number;

Step 2: Construct *(N-m+1) m*-dimensional equally spaced vectors from the initial time-series data as:

$$X_i^m=\left[u(i),u(i+1),...,u(i+m-1)\right]\ ,i=1,2,...,N-m+1 \tag{4}$$

Step 3: For each X_i^m, compute the measure that describes the similarity between it and the others as:

$$X_j^m \text{ as: } C_i^m(r)=\frac{Number\ of\ \ j\ such\ that\ d_{ij}\le r}{N-m+1} \tag{5}$$

Where
$$d_{ij}=\max_{k=1,2,...,m}\left(\left|X_i^m(k)-X_j^m(k)\right|\right)$$
$$=\max_{k=1,2,...,m}\left(\left|u(i+k-1)-u(j+k-1)\right|\right).$$

Step 4: Define the following function:

$$\Phi^m(r)=\frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\ln\left(C_i^m(r)\right) \tag{6}$$

Step 5: Repeat the previous steps 2 to 4 and compute $\Phi^{m+1}(r)$.

Step 6: Then, the ApEn is defined by:

$$ApEn(m,r,N)=\Phi^m(r)-\Phi^{m+1}(r) \tag{7}$$

2.3.2. Sample Entropy (SampEn)

Due to the fact that ApEn inherently includes a bias towards regularity, as it will count a self-match of vectors, Richman and Moorman [18] develop a sample entropy (SampEn) that does not count a self-match and thus eliminating the bias towards regularity. In addition to eliminating self-matches, the SampEn algorithm is simpler than the ApEn algorithm.

For a given time-series data $\{u(i), 1\le i\le N\}$ of length N, SampEn is depicted as follows.

Step 1: Fix the embedding dimension m as a positive integer, and the relative tolerance limit r as a positive real number;

Step 2: Construct *(N-m+1) m*-dimensional and *(N-m) (m+1)*-dimensional vectors from the initial time-series data using equations (8) and (9), respectively:

$$X_i^m=\left[u(i),u(i+1),...,u(i+m-1)\right]\ ,?\ i=1,2,...\ ,N-m+1 \tag{8}$$

$$Y_i^m=\left[u(i),u(i+1),...,u(i+m)\right]\ ,?\ i=1,2,...\ ,N-m \tag{9}$$

Step 3: Using only the first *(N-m) m*-dimensional vectors, and all the *(N-m) (m+1)*-dimensional vectors: for each X_i^m and Y_i^m, compute the similarity measure between it and its others $X_{j\ne i}^m$ and $Y_{j\ne i}^m$, respectively as:

$$A_i^m(r)=\frac{Number\ of\ \ j\ such\ that\ d_{ij}^X\le r}{N-m-1} \tag{10}$$

$$B_i^m(r)=\frac{Number\ of\ \ j\ such\ that\ d_{ij}^Y\le r}{N-m-1} \tag{11}$$

Where
$$d_{ij}^X=\max_{k=1,2,...,m}\left(\left|X_i^m(k)-X_j^m(k)\right|\right)$$
$$=\max_{k=1,2,...,m}\left(\left|u(i+k-1)-u(j+k-1)\right|\right)$$
and
$$d_{ij}^Y=\max_{k=1,2,...,m+1}\left(\left|Y_i^m(k)-Y_j^m(k)\right|\right)$$
$$=\max_{k=1,2,...,m+1}\left(\left|u(i+k-1)-u(j+k-1)\right|\right).$$

Step 4: Define the functions:

$$A^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} A_i^m(r) \quad (12)$$

$$B^m(r) = \frac{1}{N-m}\sum_{i=1}^{N-m} B_i^m(r) \quad (13)$$

Step 5: Then, the SampEn is defined as:

$$SampEn(m,r,N) = \ln\left(\frac{A^m(r)}{B^m(r)}\right) \quad (14)$$

2.3.3. Permutation Entropy (PermEn)

Different from ApEn and SampEn, Bandt and Pompe propose permutation entropy (PermEn), a parameter of average entropy, to describe the complexity of a time series [19]. The PermEn takes into account the temporal order of the values in a time series. Thus, it is robust under non-linear distortion of the signal and is also computationally efficient [20]. These advantages make it suitable for analyzing data sets of huge sizes without any preprocessing and fine-tuning of parameters.

Given a time-series data $\{u(i), 1 \le i \le N\}$ of length *N*, the necessary steps in the computation of PermEn are provided as follows.

Step 1: Fix the vector window length or embedded dimension *m* and the time delay τ as positive integers;

Step 2: Define the set of ordinal patterns $\Omega_m = \{\Pi_i, 1 \le i \le m!\}$ as the possible permutations of *{1, 2,..., m}*. For example, $\Omega_m = [\{1,2,3\};\{1,3,2\};\{2,1,3\};\{2,3,1\};\{3,1,2\};\{3,2,1\}]$ for *m=3*.

Step 3: Construct $(N-(m-1)*\tau)$ *m*-dimensional vectors from the initial time-series data:

$$V_i = [v(1), v(2), ..., v(m)] = [u(i), u(i+\tau), ..., u(i+(m-1)*\tau)] \text{ ,? } i = 1, 2, ..., N-(m-1)*\tau \quad (15)$$

Step 4: Arrange each vector V_i to an increasing order such that $V_i = [v(j_1), v(j_2), ..., v(j_m)]$. If there exist two or more elements of V_i that have the same value, their original positions can be sorted according to their appearance. Thus, the corresponding ordinal pattern of V_i is $\pi_i = \{j_1, j_2, ..., j_m\}$;

Step 5: For each ordinal pattern Π_i, the corresponding probability distribution is defined by

$$p_i = \frac{Number\ of\ occcurences\ of\ the\ ordinal\ pattern\ \Pi_i\ on\ \pi = [\pi_1;\pi_2;...;\pi_{N-(m-1)*\tau}]}{(N-(m-1)*\tau)} \quad (16)$$

Step 6: Then, the PermEn is evaluated using the Shannon entropy as:

$$PermEn(m,\tau,N) = -\sum_{i=1}^{N-(m-1)*\tau} p_i \ln(p_i) \quad (17)$$

2.3.4. Fuzzy Entropy (FuzzyEn)

In 2007, Chen et al. [21] develop a new related family of statistics called fuzzy entropy (FuzzyEn) that is freer parameter selections and more robustness to noise. The FuzzyEn as a measure of complexity excludes self-matches like SampEn.

For a time-series data $\{u(i), 1 \le i \le N\}$ of length *N*, the necessary steps in the computation of FuzzyEn are provided as follows.

Step 1: Fix the embedded dimension *m* and the parameter gradient of the boundary *n* as positive integers, and the relative tolerance limit *r* as a positive real number;

Step 2: Construct *(N-m+1)* *m*-dimensional and *(N-m)* *(m+1)*-dimensional vectors from the initial time-series data using equations (18) and (19), respectively:

$$x_i^m = [u(i), u(i+1), ..., u(i+m-1)] \text{ , } i = 1, 2, ..., N-m+1 \quad (18)$$

$$y_i^m = [u(i), u(i+1), ..., u(i+m)] \text{ , } i = 1, 2, ..., N-m \quad (19)$$

Step 3: Remove the baseline on each *m*- and *(m+1)*-dimensional vectors in order to define centered vectors:

$$X_i^m = x_i^m - \frac{1}{m}x_i^m \quad (20)$$

$$Y_i^m = y_i^m - \frac{1}{m+1}y_i^m \quad (21)$$

Step 4: Using only the first $(N-m)$ *m*-dimensional vectors, and all the $(N-m)$ *(m+1)*-dimensional vectors: for each X_i^m and Y_i^m, compute the similarity degrees Φ_i^X and Φ_i^Y between each and corresponding neighbors $X_{j\neq i}^m$ and $Y_{j\neq i}^m$, respectively as:

$$\Phi_i^X(r) = \frac{1}{N-m-1}\sum_{j=1, j\neq i}^{N-m} \mu\left(d_{ij}^X, r\right) \quad (22)$$

$$\Phi_i^Y(r) = \frac{1}{N-m-1}\sum_{j=1, j\neq i}^{N-m} \mu\left(d_{ij}^Y, r\right) \quad (23)$$

Where $d_{ij}^{X} = \max_{k=1,2,...,m}\left(\left| X_i^m(k) - X_j^m(k) \right| \right)$ and $d_{ij}^{Y} = \max_{k=1,2,...,m}\left(\left| Y_i^m(k) - Y_j^m(k) \right| \right)$. For the sake of convenience and to catch as much detailed information as possible using the fuzzy function $\mu\left(d_{ij}, r\right) = \exp\left(-(d_{ij}/r)^n\right)$, Chen et al. [22] recommend small integers greater than one of the parameter gradient of the boundary *n*.

Step 5: Define the functions:

$$\varphi^X(r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\Phi_i^X(r) \tag{24}$$

$$\varphi^Y(r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\Phi_i^Y(r) \tag{25}$$

Step 6: Then, the FuzzyEn is defined as:

$$FuzzyEn(m,n,r,N) = \ln\left(\frac{\varphi^X(r)}{\varphi^Y(r)}\right) \tag{26}$$

2.3.5. Increment Entropy (IncrEn)

Increment entropy (IncrEn) is a new approach to measure the complexity of time series introduces in 2007 by Xiaofeng et al. [23]. IncrEn indicates hidden characteristics of the dynamic changes in a signal since it can detect either structural or energetic changes because. During the process, each increment is mapped using two letters word.

Considering a time-series data $\{u(i), 1 \le i \le N\}$ of length *N*, the IncrEn algorithm is described as follows.

Step 1: Fix the window or run length $m \le N$ as a positive integer, and the resolution level *R* as a positive real number;

Step 2: Construct an increment time-series $\{v(i), 1 \le i \le N-1\}$ where $v(i) = u(i+1) - u(i)$.

Step 3: Construct *(N-m) m*-dimensional vectors from the increment time-series as

$$V_i = [v(i), v(i+1), ..., v(i+m-1)] \ , i = 1,2,...,N-m \tag{27}$$

Step 4: Map each element of the vector V_i into a word of two letters. The sign of each element is denoted as a letter (positive, negative, zero), $\delta_i(k) = \mathrm{sgn}(V_i(k)) \ , k = 1,2,...,m$. The magnitude of each element compared with other elements in the vector is quantified as another letter, $q_i(k)$, which is dependent on the quantifying resolution R. Then, V_i is mapped into a word of *2m* letters:

$$\omega_i = U_{k=1}^{m}\,\delta_i(k)q_i(k) \tag{28}$$

And for the time-series data $\{u(i), 1 \le i \le N\}$ we obtain *(N-m)* words $\{\omega_i \ , 1 \le i \le N-m\}$. In this paper, the magnitude of increment is quantified by the standard deviation of increment epoch and the quantifying resolution defined as [23],

$$q_i(k) = \begin{cases} 0 \ , std(V_i) = 0 \\ \min\left(R, \left\lfloor \dfrac{\|V_i(k)\| \times R}{std(V_i)} \right\rfloor\right) \ , std(V_i) \ne 0 \end{cases} \tag{29}$$

Where $\lfloor \ \rfloor$ round the element toward zero.

Step 5: Calculate the relative frequency of each unique word ω_n in $\{\omega_i \ , 1 \le i \le N-m\}$ as,

$$p_n = \frac{Number\ of\ occcurences\ of\ the\ unique\ word\ \omega_n\ on\ \{\omega_i \ , 1 \le i \le N-m\}}{N-m} \tag{30}$$

Step 6: Then, the IncrEn is defined using the Shannon entropy as:

$$IncrEn(m,R,N) = -\sum p_n \ln(p_n) \tag{31}$$

2.4. Kernel Machines

2.4.1. Least-Squares Support Vector Machine (LS-SVM)

Introduced by Suykens and Vandewalle [24, 25], the LS-SVM is an extension of the standard SVM that analyses data and recognizes patterns. In addition, LS-SVM is a class of kernel-based learning methods that solves linear systems. Given a training set $\{(x_k, y_k)\}_{k=1}^{N}$, with input $x_k \in IR^n$ and class label $y_k \in \{\pm 1\}$, the parameters $\omega = \sum_{k=1}^{N}\alpha_k y_k \varphi(x_k)$ and *b* of the hyperplane are obtained by solving the formulation:

$$\min_{\omega,b,e} J(\omega,b,e) = \frac{1}{2}\omega^T\omega + \frac{1}{2}\gamma\sum_{k=1}^{N}e_k^2 \tag{32}$$

Subject to the equality constraints:

$$y_k\left[\omega^T\varphi(x_k) + b\right] = 1 - e_k \quad , k = 1,2,...,N \tag{33}$$

Where γ is the regularization parameter and $e_{k,\ k=1,2,...,N}$ are the errors between the desired and the obtained outputs of the LS-SVM.

To derive the dual problem of equations (32) and (33), the Lagrange multipliers $\alpha_{k,\ k=1,2,...,N}$ are used:

$$L(\omega,b,e;\alpha)=J(\omega,b,e)-\sum_{k=1}^{N}\alpha_k\left\{y_k\left[\omega^T\varphi(x_k)+b\right]-1+e_k\right\} \tag{34}$$

And the conditions for optimality are defined by:

$$\begin{cases}\dfrac{\partial L}{\partial\omega}=0 \;\rightarrow\; \omega=\sum_{k=1}^{N}\alpha_k y_k\varphi(x_k)\\ \dfrac{\partial L}{\partial b}=0 \;\rightarrow\; \sum_{k=1}^{N}\alpha_k y_k=0\\ \dfrac{\partial L}{\partial e_k}=0 \;\rightarrow\; \alpha_k=\gamma e_k \quad ,k=1,2,...,N\\ \dfrac{\partial L}{\partial\alpha_k}=0 \;\rightarrow\; y_k\left[\omega^T\varphi(x_k)+b\right]-1+e_k=0 \quad ,k=1,2,...,N\end{cases} \tag{35}$$

Can be written as the linear system:

$$\begin{bmatrix}0 & Y^T\\ Y & \Omega+\gamma^{-1}I\end{bmatrix}\begin{bmatrix}b\\ \alpha\end{bmatrix}=\begin{bmatrix}0\\ \vec{1}\end{bmatrix} \tag{36}$$

with $Y=\left[y_1,y_2,...,y_N\right]^T$, $\vec{1}=[1,1,...,1]^T$, $e=\left[e_1,e_2,...,e_N\right]^T$, $\alpha=\left[\alpha_1,\alpha_2,...,\alpha_N\right]^T$ and I is the identity matrix. Mercer's conditions for kernels are applied within the matrix Ω with elements given as $\Omega_{kl}=y_k y_l\varphi(x_k)^T\varphi(x_l)=y_k y_l K(x_k,x_l)$.

Hence, the obtained classifier is defined by:

$$y(x)=sign\left[\omega^T\varphi(x_k)+b\right]=sign\left[\sum_{k=1}^{N}\alpha_k y_k K(x,x_k)+b\right] \tag{37}$$

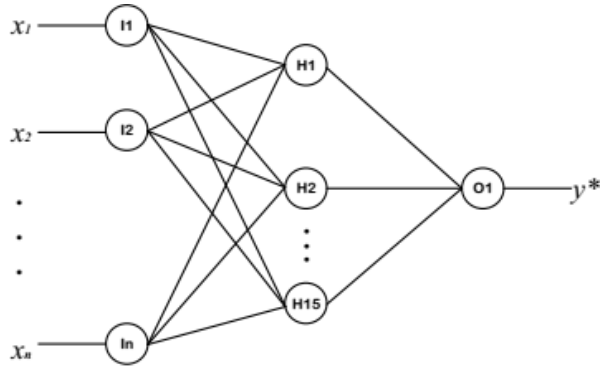

Figure 4. *sMLPNN architecture.*

2.4.2. Simple Multi-layers Perceptron Neural Network (sMLPNN)

The MLPNN is a popular kernel machine for data processing [26-29]. The goal is to discriminate F and NF EEG signals using a sMLPNN. As shown in Figure 4, the sMLPNN has *n*, fifteen and one neurons at the input, hidden and output layers, respectively. Here, we have a unique hidden layer and the input layer is not really treated as a layer of a neural processing unit. No processing will occur in the input layer and it is only an input vector augmented with a bias term, whose components will be fed to the next layer. The advantage of using this sMLPNN is the rapid execution and generalization of the trained network, which is particularly advantageous in the detection of EEG signals applications.

The weights and bias are determined using the backpropagation algorithm, which is based on searching a minimum mean square error between the desired and obtained solutions using gradient descent [27]. The backpropagation algorithm is performed repeatedly until the sMLPNN solution agrees with the desired value within a pre-specified tolerance.

2.5. Experiments and Performance Measures

The discrimination of F and NF EEG signals is closely related to the clinical application of locating epilepsy-affected brain areas. In this work, the Kruskal-Wallis test, the LS-SVM with radial basis function (RBF) kernel ($K(x,x_k)=\exp\left(-\dfrac{1}{2\sigma^2}\|x_k-x\|\right)$), the sMLPNN with fifteen neurons on the unique hidden layer, the 10-fold cross-validation technique, and the first 750 pairs of F and NF bivariate EEG signals of the Bern-Barcelona EEG dataset [14] are used for studying the performances of the proposed method.

Initially, each entropy measure (ApEn, SampEn, PermEn, FuzzyEn, or IncrEn) is computed on the EEG signal and its corresponding rhythms for the X and Y time series of each F and NF EEG signal. Furthermore, the sum of each entropy measure on the X and Y time series (X+Y) is computed. In addition, a statistical procedure used to compare several populations in terms of mean namely the Kruskal-Wallis test is performed to determine the statistical significance of the different X+Y entropy measures. The procedure returns the p-value for the null hypothesis that F and NF entropy measures were obtained from the same population (or equivalently, from different populations with the same distribution). If the p-value is near zero, this casts doubt on the null hypothesis and suggests that F and NF entropy measures are significantly different. It is common to declare a result significant if the p-

value is less than 0.05. This helps to design different experiments as follows. Each X+Y entropy measure on EEG and corresponding rhythms firstly defines a 6-dimension feature vector; secondly, all computed X+Y entropies are put together to define a 30-dimension feature vector. Thereafter, all computed X+Y entropies are used to define a 5-dimension feature vector on EEG and each rhythm.

For each experiment, the 10-fold cross-validation technique is used to determine the parameters of kernel machines. In the 10-fold cross-validation technique, the data is once permutated randomly and partitioned into 10 equal disjoint subsets. In the *i-th* (*i=1, 2,…, 10*) iteration, the *i-th* subset is used to estimate the parameters of the model trained on the other 9 subsets. Finally, the 10 different estimates parameters of the model are combined and averaged to obtain the final parameters of the model. The 10-fold scheme is used to achieve best performances which are evaluated using three measures defined as:

Sensitivity:

$$Sen(\%) = 100 \times \frac{T_P}{T_P + F_N} \tag{38}$$

Specificity:

$$Spe(\%) = 100 \times \frac{T_N}{T_N + F_P} \tag{39}$$

Accuracy:

$$Acc(\%) = 100 \times \frac{T_P + T_N}{T_P + F_N + T_N + F_P} \tag{40}$$

Where T_P, F_N, T_N and F_P are the total number of true positive (F), false negative, true negative (NF) and false positive, respectively.

3. Results and Discussion

The JPTs-based entropies described in the previous section have been applied to 750 F and 750 NF bivariate EEG signals. In this work, the parameters used to compute different entropy measures are determined empirically after a multitude of tests. At the end of the procedure, computed measures on X and Y time series are ApEn (3, 0.2, 10240), SampEn (3, 0.2, 10240), PermEn (3, 1, 10240), FuzzyEn (3, 2, 0.2, 10240) and IncrEn (3, 4, 10240). This work studies different time series EEG signals of durations corresponding to 20 s. The Kruskal-Wallis statistical test is performed to find the statistical significance ($p < 0.05$) of the computed X+Y entropy measures over EEGs and different rhythms. Table 1 presents the results of the statistical analysis of the computed X+Y entropies using the X and Y time series.

Table 1. *Statistical analysis of the proposed JPTs-based X+Y entropy measures.*

EEG or rhythms		Statistical measures	Entropy measures				
			ApEn	**SampEn**	**PermEn**	**FuzzyEn**	**IncrEn**
EEG		F: Mean (SD)	0.738 (0.220)	0.619 (0.205)	4.310 (0.285)	0.509 (0.172)	8.664 (0.605)
		NF: Mean (SD)	0.885 (0.288)	0.776 (0.273)	4.251 (0.293)	0.652 (0.232)	8.530 (0.623)
		p-value	6.49×10^{-26}	6.61×10^{-31}	0.05×10^{-0}	2.492×10^{-37}	0.02×10^{-0}
EEG rhythms separation using DLT	δ	F: Mean (SD)	0.289 (0.046)	0.267 (0.041)	1.752 (0.023)	0.131 (0.022)	2.233 (0.051)
		NF: Mean (SD)	0.279 (0.063)	0.268 (0.049)	1.760 (0.028)	0.139 (0.034)	2.253 (0.056)
		p-value	2.53×10^{-7}	2.09×10^{-2}	3.01×10^{-21}	0.05×10^{-0}	8.23×10^{-18}
	θ	F: Mean (SD)	0.644 (0.051)	0.646 (0.130)	2.088 (0.018)	0.479 (0.066)	3.030 (0.045)
		NF: Mean (SD)	0.666 (0.047)	0.647 (0.121)	2.098 (0.016)	0.500 (0.059)	3.064 (0.049)
		p-value	2.06×10^{-42}	2.87×10^{-16}	5.74×10^{-31}	1.08×10^{-40}	7.21×10^{-28}
	α	F: Mean (SD)	0.760 (0.056)	0.675 (0.071)	2.486 (0.016)	0.680 (0.081)	3.930 (0.042)
		NF: Mean (SD)	0.781 (0.038)	0.690 (0.059)	2.505 (0.021)	0.710 (0.070)	3.941 (0.051)
		p-value	2.38×10^{-5}	4.89×10^{-9}	0.96×10^{-0}	9.11×10^{-13}	0.62×10^{-0}
	β	F: Mean (SD)	0.975 (0.101)	0.905 (0.115)	3.271 (0.043)	1.055 (0.141)	5.744 (0.089)
		NF: Mean (SD)	0.998 (0.089)	0.956 (0.088)	3.250 (0.045)	1.099 (0.099)	5.689 (0.092)
		p-value	0.03×10^{-0}	5.89×10^{-10}	9.46×10^{-18}	1.21×10^{-10}	3.02×10^{-19}
	γ	F: Mean (SD)	1.534 (0.193)	1.331 (0.253)	4.372 (0.032)	1.710 (0.191)	8.130 (0.074)
		NF: Mean (SD)	1.628 (0.110)	1.389 (0.172)	4.366 (0.020)	1.799 (0.110)	8.128 (0.060)
		p-value	4.45×10^{-37}	8.84×10^{-14}	4.82×10^{-0}	2.75×10^{-40}	2.91×10^{-0}
EEG rhythms separation using DChT	δ	F: Mean (SD)	0.283 (0.047)	0.210 (0.039)	1.758 (0.021)	0.116 (0.026)	2.226 (0.046)
		NF: Mean (SD)	0.272 (0.063)	0.205 (0.041)	1.767 (0.025)	0.121 (0.033)	2.246 (0.056)
		p-value	1.07×10^{-7}	0.09×10^{-0}	1.18×10^{-21}	0.00×10^{-0}	5.60×10^{-21}
	θ	F: Mean (SD)	0.639 (0.049)	0.529 (0.060)	2.112 (0.020)	0.415 (0.049)	3.032 (0.045)
		NF: Mean (SD)	0.667 (0.044)	0.555 (0.057)	2.124 (0.020)	0.450 (0.046)	3.060 (0.045)
		p-value	2.48×10^{-41}	1.68×10^{-22}	1.03×10^{-31}	5.47×10^{-48}	5.39×10^{-30}
	α	F: Mean (SD)	0.765 (0.054)	0.670 (0.069)	2.508 (0.015)	0.671 (0.080)	3.920 (0.034)
		NF: Mean (SD)	0.778 (0.040)	0.690 (0.051)	2.507 (0.017)	0.702 (0.061)	3.916 (0.038)
		p-value	2.38×10^{-5}	3.93×10^{-10}	0.93×10^{-0}	7.04×10^{-16}	0.37×10^{-0}
	β	F: Mean (SD)	0.971 (0.111)	0.856 (0.113)	3.277 (0.040)	1.042 (0.148)	5.734 (0.083)
		NF: Mean (SD)	0.993 (0.090)	0.894 (0.080)	3.256 (0.043)	1.092 (0.099)	5.684 (0.089)
		p-value	0.00×10^{-0}	3.88×10^{-14}	7.26×10^{-18}	3.33×10^{-12}	1.82×10^{-23}

EEG or rhythms	Statistical measures	Entropy measures				
		ApEn	SampEn	PermEn	FuzzyEn	IncrEn
γ	F: Mean (SD)	1.529 (0.191)	1.214 (0.186)	4.396 (0.034)	1.647 (0.177)	8.134 (0.072)
	NF: Mean (SD)	1.629 (0.110)	1.297 (0.108)	4.394 (0.024)	1.748 (0.089)	8.123 (0.056)
	p-value	5.44×10^{-36}	7.42×10^{-20}	0.21×10^{-0}	1.52×10^{-47}	0.01×10^{-0}

Taking in account the results of the above statistical analysis, the comparison of some computed measures is projected and presented in Figure 5 and Figure 6. These projections can be used to visualize and determine if the computed X+Y entropy measures are able to discriminate F and NF EEG signals.

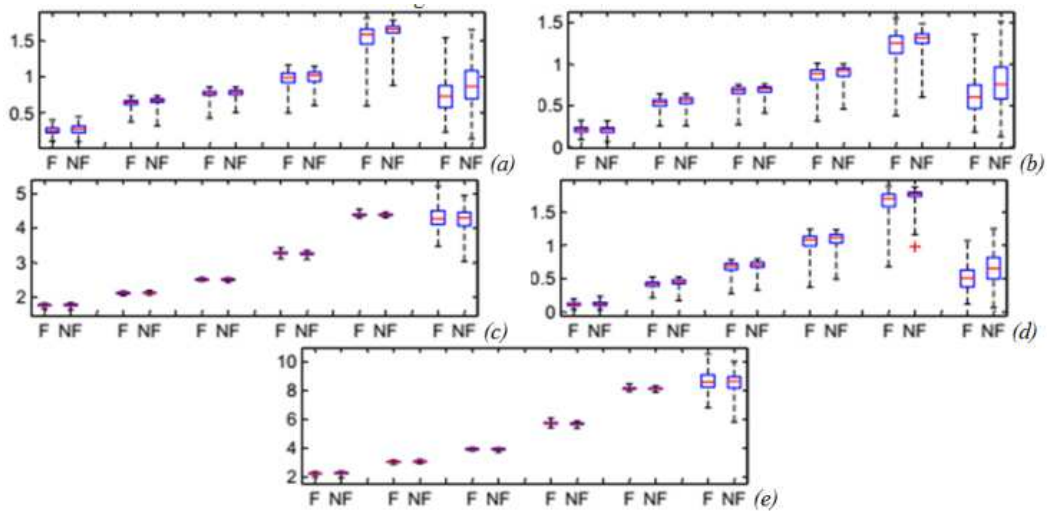

Figure 5. *Boxplot comparison of some DLT-based X+Y entropy measures. Where (a) to (e) represent the ApEn, SampEn, PermEn, FuzzyEn and IncrEn, respectively. For each projection, boxes are grouped two by two such that from the left to the right we have the delta to gamma rhythms, and EEG, respectively.*

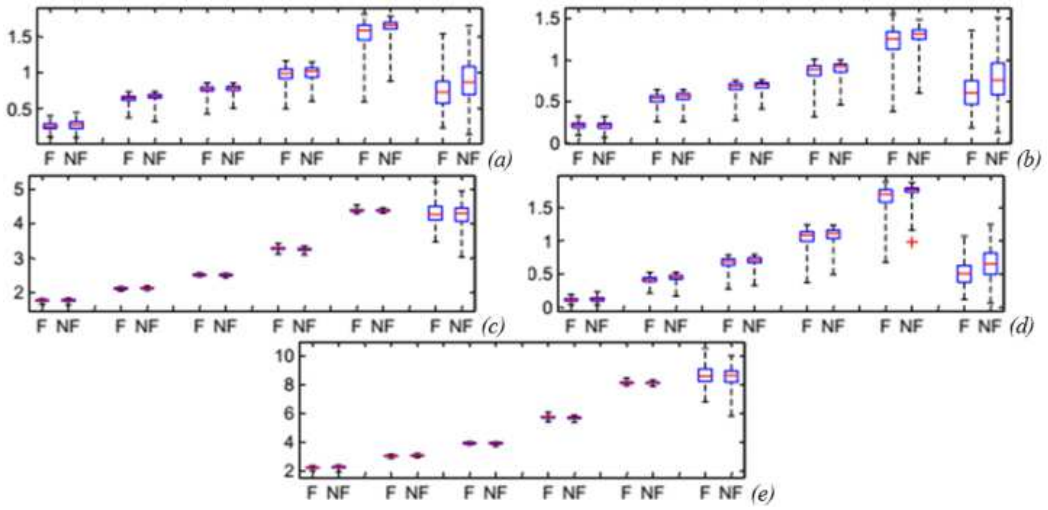

Figure 6. *Boxplot comparison of some DChT-based X+Y entropy measures. Where (a) to (e) represent the ApEn, SampEn, PermEn, FuzzyEn and IncrEn, respectively. For each projection, boxes are grouped two by two such that from the left to the right we have the delta to gamma rhythms, and EEG, respectively.*

Table 1 reports the analysis of the EEG and corresponding rhythms in terms of mean, standard deviation (SD) and statistical significance for the discrimination of F and NF EEG signals. From Table 1, Figures 5 and Figure 6, it is clearly observed that DLT and DChT extend to give similar results; and in low frequency, entropy measures computed on F EEG signals extend to be higher than the ones on NF EEG signals in EEG and EEG rhythms for analysis. This supports the outcomes of the previous study using the same database [3]. The measures in different rhythms increased with the frequency band occupation and the *p*-values of the theta rhythm are statistically more significant to discriminate F and NF EEG signals. In addition, it is also shown that for some entropy measures, F data have the lowest values, which means that the NF data are more regular, periodic and predictable than the F data. This implies that JPTs-based entropy measures can give better discrimination of F and NF EEG signals. It is also observed that neither of the entropy

measures by themselves sufficiently discriminate F and NF time series of EEG signals. In a few cases, the overlaps of the computed entropy measures are significantly high. This implies that a simple threshold applied to these measures will not be sufficient to distinguish the F and NF time series of EEG signals. Overall, this experimental analysis of entropy measures based on statistical analysis using the Kruskal-Wallis test and visual projection using boxplots demonstrated that computed JPTs-based entropy measures try to explicitly model differences between class labels within the data and successful discrimination for different experiments can be achieved using classifier machines.

Thus, keeping in mind the outcome of the above statistical analysis and the comparison on 750 F and 750 NF EEG signals using our proposed methodology, we use two kernel machines namely LS-SVM and sMLPNN to discriminate F and NF EEG signals. For different pre-defined experiments, all the first 750 F and 750 NF EEG bivariate signals of the database are considered. To find the optimal classifier machine, the first 60% of each set of bivariate signals are used as the training set and the last 40% as the testing set. The training data set and the 10-fold cross-validation technique are used to train and determine the parameters of the kernel machine, while the testing data set is used to verify the accuracy of the trained kernel machine. In this work, the LS-SVM toolbox proposed by Suykens et al. [25] and the artificial neural network (ANN) Matlab toolbox are used, and obtained results are summarized in Table 2.

Table 2. *Discrimination performances of kernel machines using JPTs-based X+Y entropy measures computed on EEG and rhythms of the X and Y time series.*

Data		**Kernel machines with corresponding discrimination performances (%)**					
		DLT-based X+Y entropy measures					
		LS-SVM			**sMLPNN**		
		Sen	**Spe**	**Acc**	**Sen**	**Spe**	**Acc**
Entropy measure on EEG and rhythms	ApEn	98.67	57.67	78.17	86.00	84.00	85.00
	SampEn	88.67	65.00	76.83	81.67	81.00	81.33
	PermEn	71.67	83.33	77.50	78.67	77.67	78.17
	FuzzyEn	81.67	75.00	78.33	84.67	86.33	85.50
	IncrEn	72.67	81.67	77.17	75.00	79.67	77.33
	All entropies	86.00	81.67	83.33	97.33	96.67	97.00
All entropies on EEG or rhythm	EEG	82.33	70.00	76.17	84.67	85.33	85.00
	Delta	88.00	62.33	75.17	83.67	81.67	82.67
	Theta	66.33	86.67	76.50	81.67	78.33	80.00
	Alpha	91.00	63.33	77.17	84.67	86.33	85.50
	Beta	72.68	81.33	77.00	81.67	77.67	79.67
	Gamma	86.00	81.67	83.33	89.00	90.33	89.67

Table 2. *Continued.*

Data		**Kernel machines with corresponding discrimination performances (%)**					
		DChT-based X+Y entropy measures					
		LS-SVM			**sMLPNN**		
		Sen	**Spe**	**Acc**	**Sen**	**Spe**	**Acc**
Entropy measure on EEG and rhythms	ApEn	84.67	77.67	81.17	88.33	82.00	85.17
	SampEn	89.33	67.00	78.17	83.67	81.67	82.67
	PermEn	72.33	82.00	77.17	82.00	87.33	84.67
	FuzzyEn	91.67	69.33	80.50	87.33	85.67	86.50
	IncrEn	89.00	67.67	78.33	83.00	86.67	84.83
	All entropies	84.67	81.33	83.00	98.33	98.00	98.17
All entropies on EEG or rhythm	EEG	83.33	65.67	74.50	84.67	85.33	85.00
	Delta	88.00	62.33	75.17	80.33	84.00	82.17
	Theta	87.00	67.67	77.33	84.67	85.67	80.17
	Alpha	91.00	63.33	77.17	87.67	84.00	85.83
	Beta	73.33	83.00	78.17	72.33	90.33	81.33
	Gamma	85.00	81.67	83.33	88.00	92.67	90.33

Table 2 presents the performances of the LS-SVM and sMLPNN classifiers using different entropy measures as inputs. In general, Table 2 demonstrates that lower performances are obtained when each entropy measure is used individually. Thus, using all measures in EEG and rhythms gives a more discriminative feature vector and our automated discrimination system extends to be more accurate than using individual measures. In addition, the feature vector defined in gamma rhythm using all entropy measures is more discriminative than the feature vector defined in EEG or other rhythms, and higher performance accuracies of 89.67% and 90.33% are obtained with sMLPNN using the DLT and DChT at the rhythms separation stage, respectively. The highest discrimination accuracy obtained with the LS-SVM is 83.33% using the DLT-based all entropy measures on EEG and rhythms, and the JPTs-based all entropy measures on gamma rhythm. Using the DChT-based all entropy measures on EEG and rhythms, sMLPNN obtained the highest discrimination accuracy of 98.17%. In sum, from our current study, it is found that for each considered experiment, the performances of the sMLPNN

classifier with sigmoid as activation function is better than the performances of the LS-SVM classifier with the RBF kernel function.

For comparison, several methodologies have been recently proposed to detect and discriminate F EEG signals, and hence locate epileptogenic foci. In previous studies using the same database as reported in Table 3, different features are extracted and fed as inputs of different classifiers. This comparison can help to highlight the useful potential of polynomial transforms-based entropy measures for epileptogenic zones identification. From the supervision of Table 3, it is firstly observed that several methods have been applied to separate EEG in their different rhythms before extracting or computing different features that were fed as inputs of different classifier machines. Some authors used Fourier-based methods [7, 11, 35], EMD [10, 31, 33, 36, 39], WT [4, 5, 31] and its extensions [6, 32, 34, 37] to separate EEG rhythms and extract different entropies [4-6, 10, 30-32] for the discrimination of F and NF EEG signals. In addition, the LS-SVM is an extremely used kernel machine. It is shown that most of the reported previous works have gained fewer discrimination performances than our proposed methodology. However, some methodologies extend to be more accurate than our proposed methodology. Jukic et al. [40] proposed a methodology of epileptic F region localization that applied the multi-scale principal component analysis (MSPCA) denoising algorithm, and the autoregressive (AR) Burg method to generate the power spectral density (PSD) of denoised EEG signals. Furthermore, PSD values were fed as inputs to five machine learning techniques that gained a higher performance accuracy of 98.90% using the Rotation Forest (RF) classifier. Sharma et al. [41] also proposed a methodology where third-order cumulant features were extracted and fed as inputs of SVM that gained a higher discrimination accuracy of 99.00%. Recently, an automated diagnosis methodology of F EEG signals that applied the Fast Walsh-Hadamard Transform (FWHT) method to analyze EEG signals in the frequency domain in terms of Hadamard coefficients is developed [42]. Using the decomposed Hadamard coefficients, ApEn, log-energy entropy (LogEn), FuzzyEn, SampEn, and PermEn are extracted and supplied to the ANN classifier, with a 10-fold cross-validation method to classify the F and NF classes. This methodology achieved a maximum discrimination accuracy of 99.50%.

In sum, five different entropy measures are computed in this paper and fed as inputs of two well-known kernel machines to discriminate F and NF EEG signals with acceptable performances. Thus, despite many proposed methodologies, this paper can provide a roadmap of polynomial transforms-based entropies to measure the complexity and discriminate F and NF biomedical signals.

Table 3. *Comparison of the F and NF discrimination performances using the same Bern-Barcelona EEG data.*

Authors	Frameworks	Number of bivariate F and NF EEG signals	Discrimination performances (%)		
			Sen	Spe	Acc
Zhu *et al.,* 2013 [30]	Delay permutation entropy (DPE) - SVM	50 F - 50 NF	-	-	84.00
Rajeev *et al.,* 2015b [4]	DWT - Entropy measures/Integrated index - LS-SVM	50 F - 50 NF	84.00	84.00	84.00
Manish *et al.,* 2016 [5]	Orthogonal wavelet filter banks - Entropies - LS-SVM	3750 F - 3750 NF	91.25	96.56	94.25
Das *et al.*, 2016 [31]	EMD-DWT - Entropies - KNN	50 F - 50 NF	90.70	88.10	89.40
Singh and Pachori, 2017 [7]	Fourier-based rhythms - Features - LS-SVM	50 F - 50 NF	-	-	89.70
Abhijit *et al.*, 2017 [6]	Tunable-Q WT (TQWT) - Multivariate sub-band fuzzy entropy - RF/LS-SVM	3750 F - 3750 NF	83.86	85.46	84.67
Vipin *et al.*, 2017 [32]	Flexible analytic WT (FAWT) method - Cross correntropy/Log energy entropy/SURE entropy - LS-SVM with RBF kernel	3750 F - 3750 NF	93.25	95.57	94.41
Rajeev *et al.*, 2018 [33]	EMD - Amplitude, precession and deformation bandwidth - LS-SVM	3750 F - 3750 NF	83.47	84.56	84.01
Abhijit *et al.*, 2018 [34]	Empirical WT - Area measures of the 2D reconstructed phase space - LS-SVM	750 F - 750 NF	81.60	83.46	82.53
Gupta *et al.*, 2018 [35]	Fourier-Bessel series expansion (FBSE) - 17 different features - LS-SVM	50 F - 50 NF	-	-	81.50
Acharya *et al.*, 2018 [9]	Various non-linear features - LS-SVM	3750 F - 3750 NF	89.97	85.89	87.93
Gupta and Pachori, 2019 [10]	EMD - Sharma-Mittal entropy - LS-SVM	3750 F - 3750 NF	85.78	80.45	83.18
Subasi *et al.*, 2019 [36]	EMD - Statstical features - Random forest (RF) classifier	3750 F - 3750 NF	95.80	96.10	95.94
Dalal *et al.*, 2019 [37]	Flexible time-frequency coverage analytic WT - Fractal dimension - Robust energy-based least square twin SVM	50 F - 50 NF	-	-	90.20
Fasil and Rajesh, 2019 [38]	Time domain - Exponential enrgy - SVM	3750 F - 3750 NF	-	-	89.00
Chen *et al.*, 2020 [39]	ARMA/EMD - Singular value - SVM	50 F - 50 NF	100	97.9	93.00
Vipin and Pachori, 2020 [11]	Fourier-Bessel series expansion based (FBSE) flexible time-frequency coverage WT - Mixture correntropy/Exponential energy - LS-SVM	3750 F - 3750 NF	95.47	96.24	95.85
Jukic *et al.*, 2020 [40]	MSPCA - AR/PSD - Rotation forest classifier	1000 F - 1000 NF	99.30	98.50	98.90
Sharma *et al.*, 2020 [41]	Third-order cumulant - SVM	3750 F - 3750 NF	99.33	98.66	99.00
Prasanna *et al.*, 2020 [42]	FWHT - Entropies (ApEn, LogEn, FuzzyEn, SampEn, PermEn) - ANN	3750 F - 3750 NF	99.70	99.30	99.50
This paper	Polynomial-based rhythms - Entropy measures (ApEn, ampEn, PermEn, FuzzyEn, IncrEn) - Kernel machines (LS-SVM, sMLPNN)	750 F and 750 NF	98.33	98.00	98.17

4. Conclusion

The discrimination of F and NF EEG signals using JPTs-based entropy measures and kernel machines is discussed in this paper. The recently proposed JPTs approach is used to separate bivariate EEG signals into their different EEG rhythms before computing entropy measures. Furthermore,

the Kruskal-Wallis statistical test is applied and demonstrated that JPTs-based X+Y entropy measures increased with the frequency band occupation and are statistically significant to discriminate F and NF EEG signals. To discriminate F and NF EEG signals in order to detect epileptogenic zones, computed X+Y measures are fed as inputs of two well-known kernel machines namely LS-SVM and sMLPNN, and obtained results demonstrated that the combination of all entropy measures on EEG and rhythms as inputs of the sMLPNN produced better discrimination performances. Overall, this paper demonstrates that entropy measures are suitable for the complexity or regularity analysis of F and NF EEG signals. In addition, the simple architecture of the MLPNN classifier with sigmoid as the activation function can be helpful to locate the epileptogenic focus before pre-surgical evaluation.

Formatting of Funding Sources

This research did not received any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Acharya U. R., Swapna G., Martis R. J., Suri J. S., *"Automated EEG Analysis of Epilepsy: A Review"*, Knowledge Based System, No. 45, pp. 147-165, 2013.

[2] Pati S., Alexopoulos A. V, *"Pharmacoresistant Epilepsy: From Pathogenesis to Current and Emerging Therapies"*, Cleve Clin. J. Med., No. 77, pp. 457-467; 2010.

[3] Rajeev Sharma, Ram Bilas Pachori, U. Rajendra Acharya, *"Application of Entropy Measures on Intrinsic Mode Functions for the Automated Identification of F Electroencephalogram Signals"*, Entropy, Vol. 17, pp. 669-691, 2015.

[4] Rajeev Sharma, Ram Bilas Pachori, U. Rajendra Acharya, *"An Integrated Index for the Identification of F Electroencephalogram Signals Using Discrete Wavelet Transform and Entropy Measures"*, Entropy, Vol. 17, No. 8, pp. 5218-5240, 2015. DOI: 10.3390/e17085218.

[5] Manish Sharma, Abhinav Dhere, Ram Bilas Pachori, U. Rajendra Acharya, *"An Automatic Detection of F EEG Signals Using New Class of Time-Frequency Localized Orthogonal Wavelet Filter Banks"*, Knowledge-Based Systems, 2016. DOI: 10.1016/j.knosys.2016.11.024.

[6] Abhijit Bhattacharyya, Ram Bilas Pachori, U. Rajendra Acharya, *"Tunable-Q Wavelet Transform Based Multivariate Sub-Band Fuzzy Entropy with Application to F EEG Signal Analysis"*, Entropy, Vol. 19, No. 99, 2017. DOI: 10.3390/e19030099.

[7] Pushpendra Singh and Ram Bilas Pachori, *"Classification of F and NF EEG Signals Using Features Derived From Fourier-Based Rhythms"*, Journal of Mechanics in Medicine and Biology, 2017. DOI: 10.1142/S0219519417400024.

[8] Arunkumar N., Ramkumar K., Venkatraman V., Enas Abdulhay, Steven Lawrence Fernandes, Seifeedine Kadry, Sophia Segal, *"Classification of F and Non F EEG Using Entropies"*, Pattern Recognition Letters, No. 000, pp. 1-6, 2017.

[9] Acharya U. Rajendra, Y. Hagiwara, S. N. Deshpande, S. Suren, J. E. W. Koh, S. L. Oh, N. Arunkumar, E. J. Ciaccio, C. M. Lim, *"Characterization of Focal EEG Signals: A Review"*, Future Generation Computer Systems, 2018.

[10] Gupta Varun and Ram Bilas Pachori, *"A New Method for Classification of Focal and Non-Focal EEG Signals"*, Machine Intelligence and Signal Analysis, Springer, pp. 235-246, 2019.

[11] Gupta Vipin and Pachori Ram Bilas, *"Classification of Focal EEG Signals Using FBSE Based Flexible Time-Frequency Coverage Wavelet Transform"*, Biomedical Signal Processing and Control, 2020.

[12] Djoufack Nkengfack Laurent Chanel, Tchiotsop Daniel, Atangana Romain, Louis-Door Valérie, Wolf Didier, *"EEG Signals Analysis for Epileptic Seizures Detection Using Polynomial Transforms, Linear Discriminant Analysis and Support Vector Machines"*, Biomedical Signal Processing and Control, Vol. 62, 2020. https://doi.org/10.1016/j.bspc.2020.102141.

[13] Djoufack Nkengfack Laurent Chanel, Tchiotsop Daniel, Atangana Romain, Louis-Door Valérie, Wolf Didier, *"Classification of EEG Signals for Epileptic Seizures Detection and Eye States Identification Using Jacobi Polynomial Transforms-Based Measures of Complexity and Least-Square Support Vector Machine"*, Informatics in Medicine Unlocked, Vol., 2021. https://doi.org/10.1016/j.imu.2021.100536.

[14] Andrzejak R. G., Schindler K., Rummel C., *"Nonrandomness, Nonlinear Dependence and Nonstationarity of Electroencephalographic Recordings FromEpilepsy Patients"*, Physical Review E, Vol. 86, p. 046206, 2012.

[15] Pincus S., *"Approximate Entropy as a Measure of System-Complexity*". Proc. Natl. Acad. Sci. U.S.A., Vol. 88, pp. 2297-2301, 1991.

[16] Pincus S. and Huang W., *"Approximate Entropy: Statistical Properties and Applications"*, Commun. Stat. Theory Methods, Vol. 21, pp. 3061-3077, 1992.

[17] Pincus S., *"Approximate Entropy (ApEn) as a Complexity Measure"*, Chaos, Vol. 5, pp. 110-117, 1995.

[18] Richman Joshua S., Moorman Randall J., *"Physiological Time-Series Analysis Using Approximate Entropy and Sample Entropy"*, Am J Physiol Heart CircPhysiol, 278: H2039–H2049, 2000.

[19] Bandt C. and Pompe B., *"Permutation Entropy: A Natural Complexity Measure for Time Series"*, Phys. Rev. Lett., Vol. 88, pp. 174102-1–174102-4, 2002.

[20] Bandt C., *"Ordinal Time Series Analysis"*, Ecol. Modell., Vol. 182, pp. 229-238, 2005.

[21] Chen W., Wang Z., Xie H., Yu W., *"Characterization of Surface EMG Signal Based on Fuzzy Entropy"*, IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 15, No. 2, pp. 266-272, 2007.

[22] Chen Weiting, Zhuang Jun, Wangxin, Wang Zhizhong, *"Measuring of Complexity Using FuzzyEn, ApEn, and SampEn"*, Medical Engineering & Physics, Vol. 30, pp. 61-68, 2009.

[23] Xiaofeng Liu, Aimin Jiang, Ning Xu and Jianru Xue, *"Increment Entropy as a Measure of Complexity for Time Series"*, Entropy, Vol. 18, No. 22; 2016. DOI: 10.3390/e18010022.

[24] Suykens J. A., Vandewalle J.,*"Least Squares Support Vector Machine Classifiers"*, Neural Process. Lett, Vol. 9, pp. 293-300, 1999.

[25] Suykens J. A., De Brabanter J., Vandewalle J., Van Gestel T.,*"Least Squares Support Vector Machines"*, World Scientific, Vol. 4, 2002.

[26] Miller A. S., Blott B. H. and Hames T. K., *"Review of Neural Network Applications in Medical Imaging and Signal Processing",* Medical and Biological Engineering and Computing, Vol. 30, pp. 449-464, 1992.

[27] Fausett L., *"Fundamentals of Neural Networks Architectures, Algorithms, and Applications"*, Prentice Hall, Englewood Cliffs, NJ, 1994.

[28] Umut Orhan, Mahmut Hekim, Mahmut Ozer, *"EEG Signals Classification Using the K-Means Clustering and a Multilayer Perceptron Neural Network Model"*, Expert Systems with Applications, Vol. 38, pp. 13475-13481 2011.

[29] Atangana Romain, Tchiotsop Daniel, Kenne Godpromesse, Djoufack Nkengfack Laurent Chanel, *"EEG Signal Classification Using LDA and MLP Classifier"*, Health Informatics - An International Journal (HIIJ), Vol. 9, No. 1, pp. 14-32, February 2020. DOI: 10.5121/hiij.2020.9102.

[30] Zhu G., Li Y., Wen P. Paul, Wang S., Xi M., *"Epileptogenic Focus Detection in Intracranial EEG Based on Delay Permutation Entropy"*, Conference Proceedings, American Institute of Physics, Vol. 1559, pp. 31-36, 2013.

[31] Das A. B., Imamul M., Bhuiyan H., *"Discrimination and Classification of NFC and FC EEG Signals Using Entropy-Based Features in the EMD-DWT Domain"*, Biomedical Signal Processing and Control, Vol. 29, pp. 11-21, 2016.

[32] Vipin Gupta, Tanvi Priya, Abhishek Kumar Yadava, Ram Bilas Pachoria, U. Rajendra Acharya, "*Automated Detection of Focal EEG Signals Using Features Extracted from Flexible Analytic Wavelet Transform"*, Pattern Recognition Letters, Vol., No., 2017. DOI: 10.1016/j.patrec.2017.03.017.

[33] Rajeev Sharma, Varshney P., Ram Bilas Pachori, S. K. Vishvakarma, *"Automated System for Epileptic EEG Detection Using Iterative Filtering"*, IEEE Sensors Letters, Vol. 2, No. 4, pp. 1-4, 2018.

[34] Abhijit Bhattacharyya, Manish Sharma, Ram Bilas Pachori, Pradip Sircar, U. Rajendra Acharya, *"A Novel Approach for Automated Detection of Focal EEG Using Empirical Wavelet Transform"*, Neural Computing and Applications, Vol. 29, No. 8, pp. 47-57, 2018. DOI: https://doi.org/10.1007/s00521-016-2646-4.

[35] Gupta S., Krishna K. H., Pachori R. B., Tanveer M., "*Fourier-Bessel Series Expansion Based Technique for Automated Classification of Focal and Non-Focal EEG Signals*", 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1-6, 2018.

[36] Subasi A., Jukic S., Kevric J., *"Comparison of EMD, DWT and WPD for the Localization of Epileptogenic Foci Using Random Forest Classifier"*, Measurement, Vol. 146, pp. 846-855, 2019.

[37] Dalal M., Tanveer M., Pachori R. B., *"Automated Identification System for Focal EEG Signals Using Fractal Dimension of FAWT-Based Sub-bands Signals"*, Machine Intelligence and Signal Analysis; Springer: Singapore, Vol. 748, pp. 583-596, 2019.

[38] Fasil O. K., Rajesh R., *"Time-Domain Exponential Energy for Epileptic EEG Signal Classification"*, Neurosci. Lett., Vol. 694, pp. 1-8, 2019.

[39] Chen Z., Lu G., Xie Z., Shang W., *"A Unified Framework and Method for EEG-Based Early Epileptic Seizure Detection and Epilepsy Diagnosis"*, IEEE Access, Vol. 8, pp. 20080-20092, 2020.

[40] Jukic Samed, Saracevic Muzafer, Subasi Abdulhamit, Kevric Jasmin, *"Comparison of Ensemble Machine Learning Methods for Automated Classification of Focal and Non-Focal Epileptic EEG Signals"*, Mathematics, Vol. 8, No. 1481, 2020. DOI: 10.3390/math8091481.

[41] Sharma Rahul, Pradip Sircar, Ram Bilas Pachori, *"Automated Focal EEG Signal Detection Based on Third Order Cumulant Function"*, Biomedical Signal Processing and Control, Vol. 58, 2020. DOI: https://doi.org/10.1016/j.bspc.2020.101856.

[42] Prasanna J., M. S. P. Subathra, Mazin Abed Mohammed, Mashael S. Maashi, Begonya Garcia-Zapirain, N. J. Sairamya, S. Thomas George, *"Detection of Focal and Non-Focal Electroencephalogram Signals Using Fast Walsh-Hadamard Transform and Artificial Neural Network"*, Sensors, Vol. 20, No. 4952, 2020. DOI: 10.3390/s20174952.