

Automated Epileptic Seizure Detection by Analyzing Wearable EEG Signals Using Extended Correlation-Based Feature Selection

Yao Guo¹, Yuan Zhang^{1*}, Md Mursalin¹, Wenyao Xu² and Benny Lo³

Abstract—Electroencephalogram (EEG) that measures the electrical activity of the brain has been widely employed for diagnosing epilepsy which is one kind of brain abnormalities. With the advancement of low-cost wearable brain-computer interface devices, it is possible to monitor EEG for epileptic seizure detection in daily use. However, it is still challenging to develop seizure classification algorithms with a considerable higher accuracy and lower complexity. In this study, we propose a lightweight method which can reduce the number of features for a multiclass classification to identify three different seizure statuses (i.e., Healthy, Interictal and Epileptic seizure) through EEG signals with a wearable EEG sensors using Extended Correlation-Based Feature Selection (ECFS). More specifically, there are three steps in our proposed approach. Firstly, the EEG signals were segmented into five frequency bands and secondly, we extract the features while the unnecessary feature space was eliminated by developing the ECFS method. Finally, the features were fed into five different classification algorithms, including Random Forest, Support Vector Machine, Logistic Model Trees, RBF Network and Multilayer Perception. Experimental results have shown that Logistic Model Trees provides the highest accuracy of 97.6% comparing to other classifiers.

Index Terms—epileptic seizure, multi-class EEG signal, wavelet analysis, extended correlation-based feature selection

I. INTRODUCTION

Epilepsy is a serious chronic incurable brain disease. Seizures often manifested as sudden, transient movement, autonomic nervous or psychiatric symptoms, such as abnormalities, affecting patients' physical and intellectual development and it could even be life-threatening [1][2]. Up to now, EEG is the most effective manner to diagnose epilepsy [3]. Epilepsy is diagnosed through visual examination of EEG signals to find abnormalities. However, this method is not only time-consuming but also leads to misdiagnose due to subjectivity and retrospective analysis. Therefore, it is necessary to find a real-time and accurate automatic detection technology.

At present, focused on the signal complexity and relativity, the time domain, frequency domain and time-frequency

domain of EEG signals are scrutinized by the nonlinear analysis methods. In the case of epilepsy automatic detection mostly used ways are wavelet transform, entropy and some synchronization parameters.

Gler and Beyli [4] proposed the neural network with adaptive capabilities and fuzzy logic qualitative approach and they found that the wavelet of Daubechies' shows better performance than others, and the second-order Daubechies' (db2) is more suitable than the fourth-order Daubechies' (db4) and sixth-order Daubechies' (db6) for EEG signals. Kumar et al. [5] applied Discrete Wavelet Transform (DWT) and Fuzzy Approximate Entropy (fApEn), support vector machine to detect epilepsy where fApEn values of different sub-bands were used to form feature vectors and these vectors were used as inputs to classifiers. Subasi [6] presents an expert model for epileptic seizures based on Artificial Neural Networks (ANN), using the expectation maximization (EM) algorithm. The characteristics of EEG signals were extracted by DWT. M. Z. Parvez et al. [7] proposed a standard seizure detection approaches to extract features for ictal and interictal signals using various well-known transformations method, decompositions and least square support vector machine was applied to classify. However, most of them are not suitable for lightweight low-cost EEG devices due to the required complex computation.

To enable ubiquitous monitoring with the low-cost wearable EEG devices, a lightweight detection algorithm is needed. The crucial goal of this study is to develop a system with low computational cost and optimum performance. In this research, the different features were extracted using nonlinear analysis method in the five frequency sub-band which were divided by DWT. Moreover, using ECFS we can reduce the number of features having the same performance and easily shrink the computational complexity and redundancy. Finally, the EEG recordings of patients were used for validating the proposed model and the experimental results have shown that our proposed approach would be an effective solution for the diagnosis of epilepsy.

The rest of this paper is formulated as follows. Section II presents the proposed method in detail. The experimental results and discussions are analyzed in Section III. Finally, Section IV presents concluding remarks.

II. PROPOSED METHOD

Delta, theta, alpha, beta, and gamma are the five main wave types in EEG signals. In this research, EEG signals were analyzed by DWT to decompose the signal in order to extract five physiological EEG bands using five levels with db2 wavelet. Then the features were extracted from each decomposed sub-band and the decomposed EEG data

*Corresponding Author.

Yao Guo, Yuan Zhang and Md Mursalin are with Shandong Provincial Key Laboratory of Network Based Intelligent Computing, University of Jinan, Jinan, 250022, China. (e-mail: yzhang@ujn.edu.cn)

Wenyao Xu is with Department of Computer Science and Engineering, The State University of New York (SUNY), New York, USA. (e-mail: wenyaoxu@buffalo.edu)

Benny Lo is with Hamlyn Centre, Imperial College London, London, SW7 2AZ, UK. (e-mail: benny.lo@imperial.ac.uk)

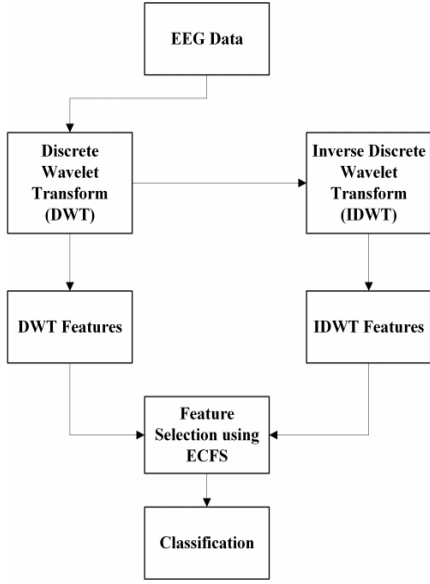


Fig.1. The block diagram for the step by step procedure of the proposed approach.

reconstructed using Inverse Discrete Wavelet Transform (IDWT) with db2 wavelet and extracted features from each reconstructed sub-band. All the extracted features were fed into the extended correlation-based feature selection (ECFS) step and then the most relevant features were selected to form the feature vector. This feature vector is used as input to the five different classifiers for classifying the EEG signal as healthy, interictal and epileptic seizure. Figure 1 describes the block diagram of the proposed approach.

A. Dataset

The EEG database used in this study was developed by the Department of Epileptology, University of Bonn, Germany [8]. Each dataset contains 100 single-channel which have 4097 data points with duration of 23.6 seconds per channel. Figure 2 illustrates the part of ideal raw EEG signals of sets H, I, and E used in this work.

B. Feature extraction

The feature extraction procedure is an essential step in processing non-stationary signals because it could be diminished the spatial dimension and advances the precision of the classification result. Three separate categories were sorted by features as follow: (1) time domain (TD); (2) frequency domain (FD); and (3) entropy based features (EB).

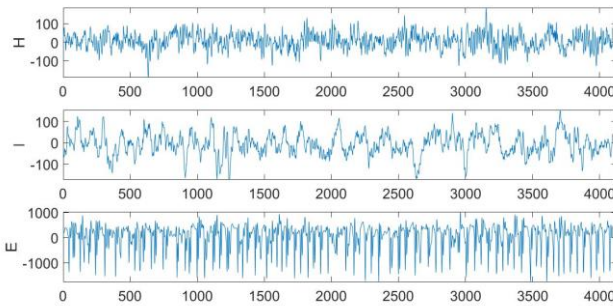


Fig.2. The ideal raw EEG signals of sets H, I, and E.

Discrete Wavelet Transform (DWT): It is very important to choose a proper wavelet and select a numeral of decomposition levels in any analysis of signals using the wavelet transform [9]. Comparing the performance of different wavelet, db2 has better property for EEG signals. EEG signals were divided into five sub-bands according to the physiological property using DWT. The wavelet coefficients were calculated for wholly five diverse sub-bands of clinical interest based on equation (1).

$$DWT(j, k) = \frac{1}{\sqrt{|2^j|}} \int_{-\infty}^{+\infty} x(t) \varphi\left(\frac{t-2^j \cdot k}{2^j}\right) dt \quad (1)$$

where a given wavelet function (t) can be scaled shifted by the scaled parameter $a_j = 2^j$ and the translation parameter $b_{j,k} = 2^j \cdot k$. Since our dataset is in range of 0-60 Hz, coefficients D1, D2, D3, D4, D5 and A5 were extracted. Matched frequencies with diverse levels of wavelet decomposition for the EEG signals were shown in TABLE I.

Inverse Discrete Wavelet Transform (IDWT): IDWT is the reverse procedure of DWT which means used to decompose the original signal while IDWT is used to assemble back into the original signal without loss of information. DWT consists of filtering and down-sampling while IDWT involves up-sampling and filtering. In case of the up-sampling process, a signal data is lengthened by adding zeros between samples. By using IDWT the original signals with a frequency below 86 Hz is assembled again. In this work, we used db2 wavelet for the signal reconstruction [10].

Extracted Feature: Feature extraction can minimize the computation cost and enhance the performance of the classifier. As the calculated wavelet coefficients show a better representation of the signals in frequency, time domain and entropy-domain, so we extracted all the features based on these and used them for feature selection and classification. Apart from the basic statistical features including standard deviation, mean and median of the wavelet coefficients in every sub-band, the entropy values denoted as a statistical measure of randomness were also calculated from the wavelet coefficients. Table II shows the extracted wavelet features. Four additional time domain features were computed to improve the performance of the classification algorithm.

TABLE I The FREQUENCY BAND SIGNAL USING 5TH LEVEL DECOMPOSITION.

Frequency Band (Hz)	Decomposed signal	Decomposition Level
43.4 - 86.8	D1	1
21.7-43.4	D2	2
10.85-32.7	D3	3
5.43-10.85	D4	4
2.71-5.43	D5	5
0-2.71	A5	5

TABLE II THE EXTRACTED FEATURES OF DIFFERENT SUB-BAND SIGNALS.

No.	Extracted features in each wavelet coefficient (WC)
1	Entropy of the WC
2	Median of the WC
3	Mean of the WC
4	Standard deviation of the WC
5	Fractal dimension of the WC

EEG signals were reconstructed using IDWT and its Fluctuation Index, Hurst Exponent, Standard Deviation, Fractal Dimension and Sample Entropy was calculated for each sub-band signals.

C. Extended Correlation-based Feature Selection (ECFS)

The ECFS is a correlation based method for feature selection which is extended with the simple statistical measurements to get features space. One of the important parts of the ECFS algorithm is a heuristic for evaluating the worth of feature subset, which is provided in equation (2),

$$r_c = \frac{kr_{cf}}{\sqrt{k+k(k-1)r_{ff}}} \quad (2)$$

where r_{cf} is the average correlation between the features and class, and r_{ff} is the average inter-correlation between all features. The foremost goal of this heuristic is to discard the irrelevant and redundant features.

To reduce the computation, information gain was applied to ranking in descending order. Features could be sorted according to relevance in equation (3). The equation (4) could be normalized to ensure the feasibility of the computation, and the symmetrical uncertainty (SU) be calculated in equation (4). The SU balances the bias for the information gain towards features with additional values ranging between 0 and 1,

$$\text{gain} = H(X) - H(X|Y) \quad (3)$$

$$\text{SU} = \frac{2 * \text{gain}}{H(Y) + H(X)} \quad (4)$$

where X and Y represent discrete values as two features and $H(X)$ and $H(Y)$ are the entropy of these two features respectively. The gain is the amount where the entropy of Y decreases showing extra information about Y given by X .

In case of ECFS, firstly we calculated a matrix of features and feature-class correlations from the given data. After that, we used the best first search technique to find out the feature subset space. We chose the subset with maximum evaluation and expanded by inserting single features. This expansion dropped back to the succeeding best subset which was not yet expanded and started from there if no improvement was observed in expanding a subset outcome. The search will stop if five consecutive non-improving subsets were found and finally the best subset returned. Then ω value was calculated for each and every selected feature by the equation (5).

$$\omega = \frac{v_{max} + v_{min}}{2} \quad (5)$$

where v_{max} is the maximum value and v_{min} is the minimum value of the data points for each feature. φ is the middle point of the ω value which is calculated by the equation (6).

$$\varphi = \frac{\omega}{2} \quad (6)$$

The φ value was then assigned to the threshold as T_{val} . If 85% of data points had the value less than the T_{val} , then we removed those features. In this way, we checked all the features with respect to T_{val} and finally got the feature set for classification. The proposed ECFS method is shown in the Algorithm 1.

Algorithm 1 ECFS algorithm for Multi-class EEG

Input: Feature Matrix

Output: Selected Features Matrix

Extract wavelet coefficient features to form a feature vector
Compute the matrix of features and feature-class correlations from a given data using Equation 2. The features are graded in descending order using r_c value.

Estimates a matrix of feature-class correlations from the training data using equation 4. The features were ranked in descending order based on their SU values

Let, *BEGIN* list contain *start state*; *END* list is empty, and *BEST* \leftarrow *start state*

while *BEGIN* is not empty **do**

i. Let, $m = \arg \max(e(x))$ (retrieve the state from *BEGIN* with maximum assessment)

ii. Discard m from *BEGIN* and insert into *END* list

if $e(m) \geq (BEST)$ **then**

BEST $\leftarrow m$

end if

iii. For each and every child node c of m which is not in the *BEGIN* or *END* list, assess and insert into the *BEGIN* list

if *BEST* altered in the previous set of expansions, **then**
go to step i

end if

iv. **return** *BEST*

end while

Compute the $T_{val} = \Phi$ value from the Equation 5 for each *BEST* value and put in the *Feature* list

for $i = 1$ to N (where the variable N indicates the number of features from *BEST* list) **do**

if $\text{Feature}[i] \leq T_{val}$; (more than 85% of data points have the value less than Φ value) **then**

discard $\text{Feature}[i]$ from the feature vector

end if

end for

III. EXPERIMENTAL RESULTS AND ANALYSIS

First, all the 500 channels of EEG data were labeled with H, I and E and data sets were decomposed into different sub-bands based on DWT of level five with db2 wavelet. The frequency ranges of different sub-bands are shown in Table I and the wavelet features shown in Table II were extracted from D1, D2, D3, D4, D5 and A5. After using the ECFS, the number of features was reduced from 49 to 19 but the accuracy has not abated too much. After that, the 5-fold cross-validation was used to separate the training and testing sets which were used as input to the five different classifiers. In Table III shows the average accuracy of five different classifiers, and Table IV shows the selected features using the ECFS. It was observed that LMT classifier shows the highest accuracy among the other classifiers. Moreover, not only the accuracy was reduced after the feature set reduction, but also greatly reduced the number of features. Table V illustrates the performance of LMT classifier in terms of three statistical measurements as precision, recall, f- measure and ROC Area. Table VI shows the comparison among different state-of-the-art methods which used the same data set for

TABLE III AVERAGE ACCURACY OF DIFFERENT CLASSIFIER.

Classifier	Accuracy using all features (%)	Accuracy after ECFS (%)
LMT	97.2	97.6
SVM	96.2	96
Random Forest	97.6	97
Multilayer Perception	96.9	96.6
RBF-Network	96.8	95.6

TABLE IV. SELECTED FEATURES USING ECFS.

No.	Feature	Details	Level
1	Wstd	Standard Deviation of detail coefficients	1,2,4,5
2	Wmean	Median of detail coefficients	6
3	Fluc	Fluctuation Index of detail coefficients	5
4	Istd	Standard Deviation of inverse discrete wavelet transform	3,5
5	Hus	Hurst Exponent of inverse discrete wavelet transform	5
6	Saen	Sample Entropy of inverse discrete wavelet transform	2,3
7	FractalDim	Fractal Dimension of inverse discrete wavelet transform	1,2,3,5

TABLE V. STATISTICAL MEASUREMENTS OF LMT CLASSIFIER.

Set	Precision	Recall	F-Measure	ROC Area
H	0.97	0.985	0.978	0.998
I	0.975	0.97	0.972	0.996
E	0.99	0.97	0.98	0.998

TABLE VI. COMPARISON BETWEEN OUR PROPOSED METHOD AND OTHER RECENT METHODS.

Authors	Methods	Accuracy (%)
Kabir et al. (2015) [11]	LMT classifier using Optimum Allocation Technique	95.33
Peker et al. (2015) [12]	Complex-valued Neural Networks (CVANN)	98.28
Acharya et al. (2012) [13]	Entropies-Fuzzy Classifier	98.10
Orhan et al. (2011) [14]	Multilayer Perceptron Neural Network (MLPNN)	95.60
Ghosh-Dastidar et al. (2008) [15]	PCA, Chaos theory and wavelet analysis, radical basis function neural network	96.73
Maheshkumar et al. (2015) [16]	Non-linear feature using least square support vector machine	82.22
Mousavi et al. (2008) [17]	wavelet decomposition, AR model, MLP classifier	96
Proposed model	Extended Correlation Feature selection and LMT classifier	97.6

multi-class problem. The experimental outcomes showed that our proposed approach could be a robust method for epileptic seizure detection.

IV. CONCLUSION

Feature selection is one of the most critical results of classification which affects the performance of the entire model. The main contribution of this paper lies on ECFS which can reduce the number of features and improve the accuracy of the classification better than the CFS method. The feature space was applied to five different classification algorithms which have high-performance rates, and among them, LMT classifier showed the highest accuracy of 97.6%. It could be inferred that this selection method would be an effective solution for the diagnosis of epilepsy.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 61572231, and in part by the Shandong Provincial Key Research & Development Project under Grant 2017GGX10141.

REFERENCES

- [1] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2013.
- [2] M. Milosević, A. Van de Vel, B. Bonroy, B. Ceulemans, L. Lagae, B. Vanrumste, and S. Van Huffel, "Automated detection of tonic-clonic seizures using 3-d accelerometry and surface electromyography in pediatric patients," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 5, pp. 1333–1341, 2016.
- [3] A. Ngugi, S. Kariuki, C. Bottomley, I. Kleinschmidt, J. Sander, and C. Newton, "Incidence of epilepsy a systematic review and metaanalysis," *Neurology*, vol. 77, no. 10, pp. 1005–1012, 2011.
- [4] I. Guler and E. D. "Ubeyli, "Adaptive neuro-fuzzy inference system " for classification of eeg signals using wavelet coefficients," *Journal of Neuroscience Methods*, vol. 148, no. 2, pp. 113–121, 2005.
- [5] Y. Kumar, M. Dewal, and R. Anand, "Epileptic seizure detection using dwt based fuzzy approximate entropy and support vector machine," *Neurocomputing*, vol. 133, pp. 271–279, 2014.
- [6] A. Subasi, "Eeg signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications*, vol. 32, no. 4, pp. 1084–1093, 2007.
- [7] M. Z. Parvez and M. Paul, "Epileptic seizure detection by analyzing EEG signals using different transformation techniques," *Neurocomputing*, vol. 145, pp. 190–200, 2014.
- [8] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Physical Review E*, vol. 64, no. 6, p. 061907, 2001.
- [9] O. Faust, U. R. Acharya, H. Adeli, and A. Adeli, "Wavelet-based eeg processing for computer-aided seizure detection and epilepsy diagnosis," *Seizure*, vol. 26, pp. 56–64, 2015.
- [10] W.-Y. Hsu, C.-H. Lin, H.-J. Hsu, P.-H. Chen, and I.-R. Chen, "Wavelet-based envelope features with automatic EOG artifact removal: Application to single-trial EEG data," *Expert Systems with Applications*, vol. 39, no. 3, pp. 2743–2749, 2012.
- [11] E. Kabir, Y. Zhang *et al.*, "Epileptic seizure detection from eeg signals using logistic model trees," *Brain Informatics*, vol. 3, no. 2, pp. 93–100, 2016.
- [12] M. Peker, B. Sen, and D. Delen, "A novel method for automated diagnosis of epilepsy using complex-valued classifiers," *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 1, pp. 108–118, 2016.
- [13] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K.-H. Ng, and J. S. Suri, "Automated diagnosis of epileptic eeg using entropies," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 401–408, 2012.
- [14] U. Orhan, M. Hekim, and M. Ozer, "EEG signals classification using thek-means clustering and a multilayer perceptron neural network model," *Expert Systems with Applications*, vol. 38, no. 10, pp. 13 475–13 481, 2011.
- [15] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 2, pp. 512–518, 2008.
- [16] M. H. Kolekar and D. P. Dash, "A nonlinear feature based epileptic seizure detection using least square support vector machine classifier," in *TENCON 2015-2015 IEEE Region 10 Conference*. IEEE, 2015, pp. 1–6.
- [17] S. Mousavi, M. Niknazar, and B. V. Vahdat, "Epileptic seizure detection using ar model on eeg signals," in *Biomedical engineering conference, 2008. CIBEC 2008. Cairo International*. IEEE, 2008, pp. 1–4.