International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 4 (2019), pp. 508-511

Epileptic EEG Signal Classification Using Multiresolution Higuchi Fractal Dimension

Achmad Rizal*, Estananto, Ratri Dwi Atmaja

School of Electrical Engineering, Telkom University, Bandung, 40257, Indonesia.

**Corresponding author*

Abstract:

Epileptic seizure could be recognized by analyzing electroencephalogram signal. For researchers, research for automatic epileptic EEG signal for seizure detection has become one of objects of interest. This study investigated the potential application of Higuchi Fractal Dimension (HFD) as a feature extraction technique for epileptic EEG signal classification. HFD calculations were performed on a series of time interval *k* value so as to produce some HFD values as features. The results showed that the proposed technique could produce accuracy up to 97.3% for three classes of epileptic EEG signal using SVM as classifier. The result showed that several HFD values could produce high accuracy for epileptic EEG signal classification without any need for any other features.

Keywords: EEG, epileptic seizure, Higuchi Fractal Dimension, SVM

1. INTRODUCTION

As stated in [1], there are complex physiological mechanisms in human body called biological signals. The important information on this complex signals can be applied to detect a person's health condition. The signals can be electrical signals that describe an electrical activity in cells in the human body. One of many biological signals of concern is electrical brain activity or called Electroencephalogram (EEG). Analyzing this EEG signal blow up the information related to our brain activities such as emotional conditions, thinking activities and even abnormalities in our brain. Epilepsy is a disorder occurred in our brain nerves, which can be detected through some specific patterns in EEG signals. EEG signal patterns that are non-linear and non-stationary or tend to be random cause a visual interpretation difficult. Signal complex analysis, therefore, is required to obtain information characteristics from EEG.

One method of complex signal analysis often used is fractal analysis. It is widely taken for analyzing biological signals such as lung sounds [1], pathological sounds [2], and heart sounds [3]. The fractal analysis provides particular pattern information namely self-similarity that occurs in biology signals that cannot be seen visually.

In the studies of EEG seizures, many methods have been proposed. One characteristic that has been used as a feature is the signal complexity of the EEG signal. Some researchers use entropy and fractal dimensions to extract the features of EEG signals. Nicolaou and Georgiou used approximate entropy (APEN) as a feature for classifying EEG epileptic seizure [4]. The APEN was attached to the non-overlap segmented EEG signal for one second and SVM was used as the classifier. Here, linear SVM produced the highest accuracy of 93.55% in two data classes. Wavelet entropy calculated at several decomposition levels yielded an accuracy up to 93.4% for three classes of EEG data [5]. The classification of epileptic seizure EEG using FD was carried out by Schneider et al. [6]. The EEG signal was windowed using Hanning window 347 samples (at 2 seconds) with a 50% overlap. EEG Higuchi FD, Katz FD, and Sevcik FD were calculated for each signal segment and SVM was used as a classifier. The resulting accuracy reached 100% for two data classes (normal and seizure).

In this study, we used the Higuchi Fractal Dimension (HFD) as a feature of EEG signals for the EEG classification; normal, ictal, and interictal signals. The HFD used as a feature in this study was calculated at various *k* time intervals resulting in several different HFD values. For classification, we used support vector machine (SVM) with several kernels. Some of these HFD values were expected to be used to differentiate normal, ictal, and interictal EEG signals without any need for other features with reasonably high accuracy. The proposed method is expected to provide an automatic EEG signal analysis method with a simple computation.

2. MATERIAL AND METHODS

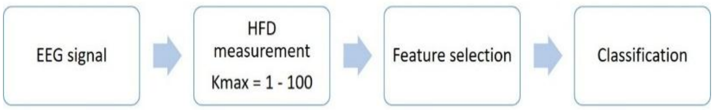

Fig 1. Diagram block of proposed method

Our proposed method is as displayed in Figure 1. We calculated HFD of each EEG signal using several time interval *k*. Then we used HFDs as feature for EEG signal. Using several feature selection method them we classified the features using SVMs. The detailed process will be explained in next subsection.

2.1 EEG Data

The EEG signal data used consisted of three classes: normal, interictal, and epileptic seizure. Normal EEG was taken from some healthy patients in a closed eye condition. Meanwhile, interictal and ictal data were recorded from five patients with long-term EEG crumbs, which were then broken down for seizure and interictal conditions. EEG signal data were recorded using a sampling frequency of 176.61 Hz with a length of each data of 23.6 s. Each data was clean of noise and filtered with LPF at a frequency of 40 Hz. Each data class consisted of 100 data so that the total data was 300. The EEG signal dataset can be accessed at http:// epileptologie-bonn.de/cms/upload/workgroup/lehnertz/ eegdata.html [7]. The examples of EEG signals in this study are presented in Figure 2.

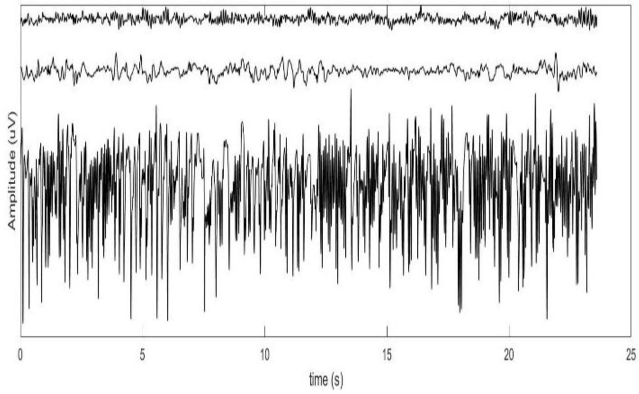

Fig 2. EEG signal, top: normal EEG, middle: interictal, below: Seizure

2.2 Higuchi Fractal Dimension

As one of the fractal dimensional measurement, the Higuchi method is often used in biomedical signals [8]. The advantages of the Higuchi method are related to its high accuracy and efficiency in measuring the fractal dimensions. If given a signal with the number of samples N, a number of lines along the *k* can be formed with different resolutions as written in Equation (1).

$$X_m^k : x(m), x(m+k), x(m+2k), \ldots, x\left(m + \left[\frac{N-m}{k}\right]k\right) \tag{1}$$

The value of *m* in Equation (1) states the initial time indication (m = 1, 2,..., k). Next, the length of the curve is explained as follows:

$$l_m(k) = \frac{\sum_{i=k}^{\lfloor N-m/k \rfloor} |x(m+ik) - x(m+(i-1)k)|(N-1)}{(\lfloor N-m/k \rfloor)k} \tag{2}$$

Notation $\lfloor a \rfloor$ means floor (a), which is a normalization factor. From Equation (2), the length of the curve could be calculated for each *k* interval as presented in Equation (3).

$$L(k) = \sum_{m=1}^{k} l_m(k) \tag{3}$$

The FD was obtained from the slope of the plot *ln(L(k))* against *ln(1/k)*, where *L(k)* is the length of curve for each interval *k* . FD values were obtained from the relationship $L(k) \alpha k^{-D}$, where the fractal dimension Higuchi (HFD) = D.

The HFD value was determined by the parameter time interval *k*. In this study, we used *k* in a certain range of values that produced the HFD value in a certain range as well. This row of HFD values will be used as a feature.

2.3 Support Vector Machine (SVM)

SVM is linear classifications, which are then developed to solve the classification of non-linear problems. Its basic concept is to get the best hyperplane for separating two data classes [9]. The hyperplane is a straight line or field that separates data between classes. The best hyperplane is obtained by maximizing the margin between two objects of different classes. The margin is the distance between the hyperplane and the closest pattern in each data class. The nearest position between the patterns of each class is called as support vector. A non-linear problem is solved by using a kernel. The method uses nonlinear fields to separate data between classes, which may not be separated by straight lines. In this study, Quadratic SVM and Cubic SVM were used as a comparison of linear SVM.

Because SVM is a method that requires a supervised learning, in this study N-fold cross-validation (NFCV) was used to separate training data and test data. In NFCV, each data class was clustered into N data sets. An N-1 data set was for the training data, and one data set was for the training data. The process was repeated up to N times so that each data set has been a test data [10]. Accuracy was taken from the average of all trials conducted [11]. The advantage of this method compared to the random distribution of training data and test data is the deviation of the lower accuracy value. The performance parameter of the proposed technique was accuracy, namely the amount of data correctly classified by the system

RESULTS AND DISCUSSION

Figure 3 shows the average HFD values for the three data classes. The HFD value increased when the *k* value increased. This was caused by the greater the *k* value, the more increasing the slope between *ln(L(k))* and *ln (1/k)* . For *k* > 10, HFD values of each data class was separated from each other so that the resulting accuracy would be quite high. Using the analysis of variance (ANOVA), we obtained F-value = 257.32 - 21,588.86 for HFD in *k* = 1-10 to *k* = 1-100. The resulting F-value was higher than F-crit (2.99).

International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 4 (2019), pp. 508-511
© International Research Publication House. http://www.irphouse.com

Table 1. Accuracy (%) using 5fold cross-validation

Classifier	Number of HFD									
	100	90	80	70	60	50	40	30	20	10
Linear SVM	96	96	96	95.7	96	96.3	97	97	97.3	86
Quadratic SVM	97	97	97.3	97.3	97	97	97	96.7	96.7	93.7
Cubic SVM	95.3	96.3	96	96.3	96	95.7	95	95	96	93.7

Table 2. Accuracy (%) using 5fold cross-validation

Classifier	Number of HFD			
	5	10	15	20
Linear SVM	76.6	77.4	79.4	79
Quadratic SVM	79.4	85.8	85.6	87.2
Cubic SVM	84.2	86.4	88.4	86.4

Table 3. Accuracy (%) using 5fold cross-validation with best first wrapper subset selection

Classifier	Accuracy
Linear SVM	96.7
Quadratic SVM	98
Cubic SVM	97.3

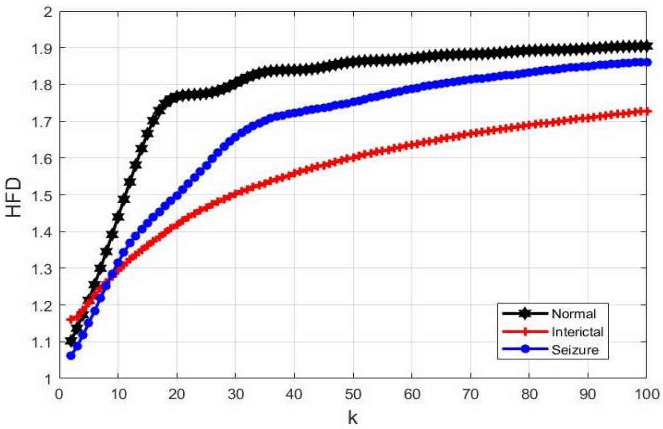

Fig 3. HFD using k = 2-100 for normal, interictal, and seizure EEG

We evaluated the classification accuracy using SVM with three kernels: linear SVM, quadratic SVM, and cubic SVM as shown in Table 1. We achieved the highest accuracy of 97.93% using 70 features and quadratic SVM. The same accuracy was also produced using 80 features but the number of features 70 was chosen as the best result for the fewer number of features.

Because the number of features needed was still quite large, the principal component analysis (PCA) algorithm was applied to reduce the number of features used from 100 to 5, 10, 15, and 20. The accuracy is shown in Table 2. The highest accuracy achieved was 88.4% using 15 PCs and cubic SVM. The result showed that the features reduction used was inappropriate for this case.

To see the possibility of increasing accuracy with fewer features, a feature subset selection (FSS) method was carried out using the wrapper method with best first search [12]. Using FSS method, we had HFD with k = 1, 2, 8, 17, 21, 26, and 73 as features. The accuracy using the three types of SVM is shown in Table 3. It then obtained the accuracy of 98% using quadratic SVM and seven features of k as previously mentioned. This result was better than the results in Table 1 using 70 features with an accuracy of 97.3%. The features selection using a subset selection wrapper was superior in terms of the number of features and higher accuracy.

The proposed method was capable of producing the high classification accuracy of EEG signals. HFD in several k values produced a number of fractal properties that could be used to distinguish normal and pathological EEGs. Compared with the method in previous studies, the proposed method produced higher accuracy with a few number of features. Table 4 shows a comparison of the methods proposed with previous research. Based on the number of features, the proposed method was inferior to the multilevel wavelet packet entropy method used by Wijayanto et al. [5], but the proposed method produced higher accuracy. Meanwhile, based on the accuracy, the method used by Schneider et al. produced 100% accuracy, but it used two data classes: seizure and normal [6]. The proposed method overall could produce higher accuracy with more classes. The use of different resolution HFDs can show some differences between normal, interictal, and ictal/seizure EEG signals. This method is different from the one used by Rizal et al. for lung sound analysis [1]. In the study, HFD was calculated on signals that have gone through a multi-scale process called the coarse-grained procedure. Meanwhile, the HFD used was calculated at only one k value. The comparison of the two methods has not been made directly because the multi-scale fractal dimension method has not been tested on EEG data. This method is open to be applied to other biomedical signals such as ECG signals, respiratory sounds or heart sounds.

Table 4. Comparison with other research

Reference	Data classes	Method	Number of Features	Classifier	Accuracy
[5]	Normal, interictal, seizure (3 classes)	Multilevel wavelet packet entropy	5	SVM	94.3%
[13]	Normal, interictal, seizure (3 classes)	Multi-distance signal level difference sample entropy	20	SVM	97.7%
[14]	Normal, interictal, seizure (3 classes)	EMD and entropy	8	SVM	97.3%
[6]	Normal, seizure (2 classes)	FD on windowed signal	23	SVM	100%
[4]	Normal, interictal, seizure (3 classes)	Approximate entropy on windowed signal	23	SVM	93.55%
[15]	Normal, interictal, seizure (3 classes)	Spectral features of EMD	7	C4.5	95.33
Proposed method without FSS	Normal, interictal, seizure (3 classes)	Multi time-interval HFD	70	SVM	97.3%
Proposed method with FSS	Normal, interictal, seizure (3 classes)	Multi time-interval HFD	7	SVM	98%

CONCLUSION

In this proposed technique, Higuchi Fractal Dimension (HFD) is applied to classify EEG signals; normal, interictal, and ictal. The HFD was calculated at several time intervals so that a series of HFD values was used as a feature. Initially, k= 2 -100 was used so that 100 HFD values were produced as features. Linear, cubic, and quadratic SVM were used to calculate the accuracy. Accuracy was calculated by reducing the number of HFD to determine the number of features that produced the highest accuracy. The highest accuracy of 97.3% was produced using 80 and 70 features with quadratic SVN as a classifier. The accuracy obtained could still be improved using the wrapper feature subset selection. As a result of wrapper FSS, using HFD on k = 1, 2, 8, 17, 21, 26, and 73, we produced the accuracy of 98%. The advantage of the proposed technique was that we used only the HFD as a feature without requiring other features. Meanwhile, the weakness of this method was that we still needed a testing to determine the number of HFDs that could produce the best accuracy. In advance, the proposed technique can be utilized for other biological signal analysis.

REFERENCES

[1] A. Rizal, R. Hidayat, and H. A. Nugroho, "Fractal Dimension for Lung Sound Classification in Multiscale Scheme," *J. Comput. Sci.*, vol. 14, no. 8, pp. 1081–1096, 2018.

[2] Z. Ali, I. Elamvazuthi, M. Alsulaiman, and G. Muhammad, "Detection of Voice Pathology using Fractal Dimension in a Multiresolution Analysis of Normal and Disordered Speech Signals," *J. Med. Syst.*, vol. 40, no. 1, pp. 1–10, 2016.

[3] M. P. Paskaš, A. M. Gavrovska, and N. B. Reljin, "Identification of Fundamental Heart Sounds from PCG Using Blanket Fractal Dimension," in *Cardiovascular Oscillations (ESGCO), 2014 8th Conference of the European Study Group on*, 2014, no. Esgco, pp. 123–124.

[4] N. Nicolaou and J. Georgiou, "Detection of epileptic electroencephalogram based on Permutation Entropy and Support Vector Machines," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 202–209, 2012.

[5] I. Wijayanto, A. Rizal, and S. Hadiyoso, "Multilevel Wavelet Packet Entropy and Support Vector Machine for Epileptic EEG Classification," in *2018 4th International Conference on Science and Technology (ICST)*, 2018, pp. 9–14.

[6] M. Schneider, P. N. Mustaro, and C. A. M. Lima, "Automatic recognition of epileptic seizure in EEG via support vector machine and dimension fractal," in *2009 International Joint Conference on Neural Networks*, 2009, pp. 2841–2845.

[7] 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," *Phys. Rev. E*, vol. 64, no. 6, p. 061907, Nov. 2001.

[8] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Phys. D Nonlinear Phenom.*,

vol. 31, no. 2, pp. 277–83, 1988.

[9] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[10] A. Rizal, R. Hidayat, and H. A. Nugroho, “Multiresolution modified grey level difference for respiratory sound classification,” *Adv. Sci. Lett.*, vol. 23, no. 5, 2017.

[11] R. Palaniappan, *Biological Signal Analysis*. Ventus Publishing ApS, 2010.

[12] R. Kohavi and G. H. John, “Wrappers for feature subset selection,” *Artif. Intell.*, vol. 97, no. 1–2, pp. 273–324, 1997.

[13] A. Rizal and S. Hadiyoso, “Sample Entropy on Multidistance Signal Level Difference for Epileptic EEG Classification,” *Sci. World J.*, vol. 2018, pp. 1–6, Sep. 2018.

[14] I. Wijayanto and A. Rizal, “Epileptic Seizure Detection in EEG Signal using EMD and Entropy,” in *the International Conference on Advancement in Data Science, E-learning and Information Systems 2019 ICADEIS2019)*, 2019.

[15] R. J. Martis *et al.*, “Application of Empirical Mode Decomposition (Emd) for Automated Detection of Epilepsy Using Eeg Signals,” *Int. J. Neural Syst.*, vol. 22, no. 06, p. 1250027, 2012.