Journal of Theoretical and Applied Information Technology
28th February 2021. Vol.99. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

EPILEPTIC SEIZURE DETECTION USING MULTIDISTANCE SIGNAL LEVEL DIFFERENCE FRACTAL DIMENSION AND SUPPORT VECTOR MACHINE

[1]DESRI KRISTINA SILALAHI, [2]ACHMAD RIZAL, [3]DIEN RAHMAWATI, [4]BANDIYAH SRI APRILLIA

[1,2,3,4]School of Electrical Engineering, Telkom University, Indonesia

E-mail: [1]desrikristina@telkomuniversity.ac.id, [2]achmadrizal@telkomuniversity.ac.id, [3]dienrahmawati@telkomuniversity.ac.id, [4]bandiyah@telkomuniversity.ac.id

ABSTRACT

Electroencephalogram (EEG) signal is a biological signal produced from an electrical activity from the brain. Abnormalities that occur in the pattern or content of the EEG signal indicate a brain disorder or disease. One of the disorders or diseases associated with brain function is epilepsy. Various methods were developed by researchers to analyze abnormalities of EEG signals using digital signal techniques. Many algorithms have been applied to achieve high performance for the classification of EEG epilepsy. However, the complexity and randomness of EEG signals is a challenge for researchers to apply the appropriate algorithm. In this research, fractal analysis of EEG signals is expected to be able to distinguish EEG signals in seizure, no seizure, and normal conditions. The multiscale method used is a multi-distance signal level difference (MSLD) and combined with the fractal dimension. Furthermore, classification is done using Quadratic SVM through 5-fold cross-validation which produces an accuracy of 99% on a scale of 1-10.

Keywords: *Classification, Epileptic Seizure, Fractal Dimension, Multi-Distance Signal Level Difference, Support Vector Machine*

1. INTRODUCTION

Electroencephalogram (EEG) signals are biological signals that result from the activity of electricity from the brain [1]. Abnormalities that occur in the pattern or content of the EEG signal indicate a brain disorder or disease. Some disorders/diseases associated with the brain function such as epilepsy [2], Alzheimer's [3], dementia [4], and abnormalities in the congenital system [5]. This disorder can be seen through changes in pattern, spectrum content and EEG signal fluctuations.

Various methods were developed by researchers to analyze EEG signal abnormalities using digital signal techniques, including wavelet analysis for epileptic EEG signal processing [6] [7], EEG signal processing for Alzheimer's cases [8], and dementia cases [9]. The whole method provides a technique for analyzing EEG signals automatically using a computer.

One of the characteristics of EEG signals that can be explored is fractal properties. This property implies the existence of self-similarity in signals that occur at different scales. In normal and pathological EEG signals there is a pattern change due to electrical process abnormalities that occur in the brain. This can change the fractal pattern of the signal. Thus the dimensional changes in fractal signals can be used to distinguish normal and pathological EEG signals [10].

Another characteristic of biological signals is multiscale properties [11]. This property is similar to fractal properties, where biological signals have similar properties on different scales. In previous studies, multiscale entropy methods were widely used for the analysis of biological signals such as electrocardiogram (ECG) signals [11], pulmonary sound signals [12], blood flow [13], and soon. The multiscale analysis will strengthen information that can be extracted from the signal so that it will be better able to distinguish the characteristics of the signal.

In this research using multiscale fractal analysis for the classification of epileptic EEG signals. This fractal analysis of EEG signals is expected to be able to distinguish EEG signals from seizure, interictal and normal EEG conditions. The multiscale method used is a multi-distance signal level difference (MSLD) [14]. The multiscale method was shown to provide high accuracy in previous studies [15]. Meanwhile, the fractal dimensions used are Higuchi fractal dimension (HFD), Katz fractal dimension (KFD), Sevcik fractal dimension (SFD), Variance fractal dimension (VFD), Petrosian C (PetC), and Petrosian D (PetD) as used in previous studies for pulmonary sound classification [16].

Feature extraction plays an important role in pattern recognition, especially in EEG signals [15]. In this research, the normalized EEG signal data was carried out in a multiscale process using the Coarse-Grained Procedure and the Multi-distance Signal Level Difference (MSLD). The multiscale process will affect the number of scales on the accuracy. Next, the fractal dimension is used as a feature to adjust the multi-scale process. The classification process used the Support Vector Machine method with n-fold cross validation. SVM is an alternative learning machine used in solving classification problems using the concept of maximum margin [17]. In the classification, three EEG data classes were used, namely EEG signals in seizure, interictal, and normal conditions. All datasets used in this research were obtained from databases available at the University of Bonn. Data were obtained from normal subjects and epilepsy subjects with ictal and interictal conditions. In a previous research that used the SVM method for classification, it produced a high level of accuracy of 97.7% by choosing the evaluation of the 5-fold cross-validation model [15].

2. RELATED WORKS

Epilepsy is a common chronic neurological disorder. Epileptic seizures are the result of temporary and unexpected electrical disorders of the brain. Research on the detection of epilepsy disorders has been widely conducted. It is even combined with several methods for classification in data mining and machine learning.

Electroencephalogram (EEG) signals are biological signals that result from the electrical activity of the brain

Shoeb AH (2010), in the results of his research by building a feature vector that unites in one feature space, the time evolution of the spectral and spatial properties of the brain's electrical activity as input data. Training was performed for 2 or more seizures per patient and tested at 916 hours of continuous EEG from 24 patients. The Support Vector Machine algorithm with Radial basis function can detect 173 test seizures with a median detection delay of 3 seconds and a median false detection rate of 2 false detections every 24 hours period. This method yields an accuracy of 96% [18].

Guler I (2007), feature extraction is performed on input data by calculating the wavelet coefficient and the Lyapunov exponent. Next, using data consisting of five data sets: each set had 5 subjects and a duration of 2360 seconds, only one set showed seizure activity and classification was performed to differentiate EEG signals. The classification accuracy of the Support Vector Machine (SVM), probabilistic neural network (PNN), and Multilayer perceptron neural network (MLPNN) methods is 99.28, 98.05, and 93.63%, respectively [19].

3. MATERIAL AND METHODS

The research flow is depicted in the block diagram in Figure 1 below.

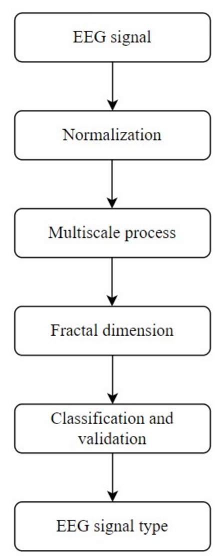


Figure 1: The Research Block Diagram

In Figure 1, the system input signal is an EEG signal consisting of three classes, namely normal, interictal, and seizure or ictal. In the EEG signal, a normalization process is carried out to equalize the signal range and remove noise. The normalization process done to homogenize the amplitude range and eliminate direct current noise in the recorded EEG signal, so that the EEG signal has an average of zero and a signal range of -1 to +1. Furthermore, a multiscale process is carried out to break the signal into several signals at different scales. The multi-scale methods are

Journal of Theoretical and Applied Information Technology
28th February 2021. Vol.99. No 4

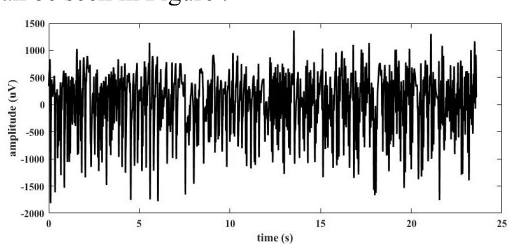
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

coarse-grained procedure and MSLD. Both methods are performed at a scale or distance (D) = 1 – 20 [11] [14] . Furthermore, the amount of these scales will be reduced to 15 and 10 to see how the number scales on the accuracy. In each signal result, the fractal dimension calculation is performed which will be used as a characteristic of the EEG signal. The fractal dimensions used in are seven fractal dimensions. The number of fractal dimensions used as a feature to adjust the scale number in a multi-scale process. The characteristics that are formed become input for the classification process to determine the class of input data. The classification used is Support Vector Machine (SVM) with various kernels. SVM requires training data and test data used n-fold cross validation for sharing training data and test data then used to determine accuracy.

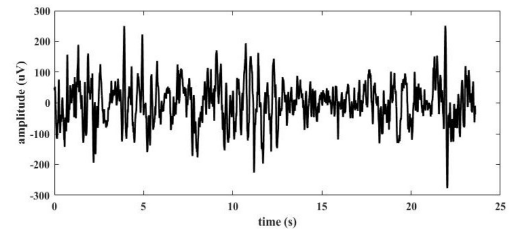
This research results can contribute to improving the classification accuracy of EEG signals in other methods by using a multi-scale process and fractal dimensions. In addition, this research results can be used as a reference in further research for the analysis of biological signals, especially EEG signals using fractal dimensions varied with other classification methods.

3.1. EEG Data

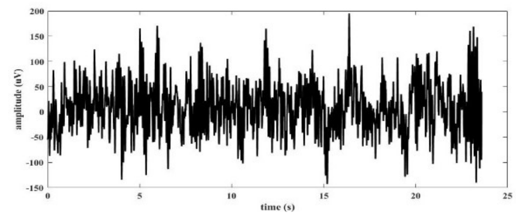
In this research, EEG data used available database at the University of Bonn [20]. Data was recorded using a 173.61 Hz sampling frequency and filtered using 40 Hz LPF. Each data has a length of 4098 samples. In this research, three classes of EEG data were used, namely EEG signals with epilepsy in conditions of seizures (ictal), EEG in patients with epilepsy when no seizures occurred (pre-ictal), and normal EEG signals in closed eye conditions. Each data class consists of 300 data. Sample data for each class can be seen in Figure .



(a)



(b)



(c)

Figure 2: (A) Ictal Condition EEG Signal (B) Pre-Ictal EEG Signal (C) Normal EEG Signal

3.2. Multidistance Signal Level Difference

Multidistance signal level difference (MSLD) is a gray-level difference (GLD) modification proposed [21]. GLD is calculated from the absolute value of the difference of two adjacent pixels in the horizontal, vertical, and diagonal directions. In the horizontal direction, GLD can be calculated as in Equation 1.

$$y(i,j)=\left|x(i,j)-x(i,j+D)\right| \quad D=\text{pixel distance} \tag{1}$$

MSLD is calculated from the difference beween the two samples at a distance of $D=1-K$, shown in the Equation 1. Equation 1 is modified into Equation 2 [14], Signal $y^d(i)$ is the output signal calculated at distance *D*. From this process, *K* new signals will be obtained.

$$y^d(i)=\left|x(i)-x(i+D)\right| \tag{2}$$

$$i=1,2,...,N-d$$

$$d=1,2,...,K$$

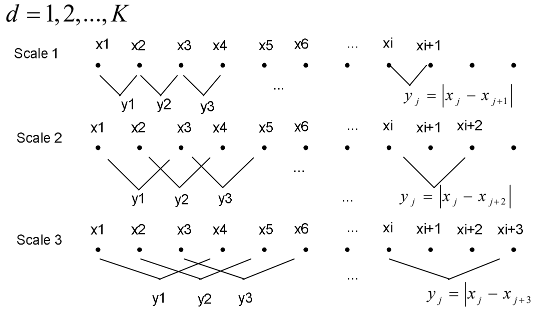


Figure 3: MSLD Illustration

Figure shows MSLD illustration. MLSD has been used to analyze other biological signals such as lung sounds [14] and EEG signals [15] with promising results. In this research, MSLD was combined with the fractal dimension for classification of EEG signals in epilepsy patients.

3.3. Fractal Dimension

One of the parameters for defining signal complexity using chaotic approaches is the fractal dimension. The fractal dimension is defined as a measure of the emergence of self-similarity, which is a signal pattern that repeats on a different scale [22]. The more signal patterns that are similar to themselves on different scales, the greater the fractal dimension value. The fractal dimension (FD) value is not an integer as in the Euclidean dimension which is 1 for lines, 2 for fields or 3 for spaces. For 1-dimensional signals, fractal dimensions are worth 1 ≤ FD <2 where the more complex a signal the value will be close to 2 [23]. Following are some measurements of FD used in this research.

3.3.1. Box Counting Method

One of the earliest fractal dimension calculation techniques is the Box Counting (BC) method. The Box counting (BC) method is motivated by the nature of the curve in filling a box-shaped space [24]. In this approach, the curve is closed by a collection of boxes, then the number of squares with a certain size is calculated to see how much is needed to cover the entire curve. At the box size close to zero, the whole curve will be closed by a box. Mathematically it can be written as in Equation 3.

$$D_B = \lim_{r\to 0}\frac{\log N(r)}{\log(1/r)} \tag{3}$$

N(r) is the number of boxes with the size r needed to cover the entire curve. Practically, the BC method estimates fractal dimensions by calculating the number of boxes needed to cover curves with various box sizes. The DB value is calculated by looking at a straight line on the log-log plot of N (r) with r. Mathematically it can be stated as in Equation 4.

$$\log N(r) = D_B \log\left(\frac{1}{r}\right) + C \tag{4}$$

The value of C is a constant, while DB is the gradient value of the graph log-log N(r) of r. This method is often called the grid method and requires a long computing time.

3.3.2. Katz Method

The Katz fractal dimension (KFD) on a curve in a row along N is defined as [25]:

$$KFD = \frac{\log_{10}(n_s)}{\log_{10}\left(\frac{d}{Lc}\right) + \log_{10}(n_s)} \tag{5}$$

KFD states fractal dimensions using the Katz method, Lc is the total length of the curve. Lc is searched by Equation 6.

$$Lc = \sum_{i=1}^{N-1} dist(i, i+1) \tag{6}$$

$dist(i,i+1)$ states the distance between two consecutive points. The d value in Equation 7 shows the farthest distance or diameter on the curve. This value can be searched by calculating the starting point of the curve with the point farthest from the starting point of the curve.

$$d = \max\left[dist(1,i)\right] \quad i = 2,\ldots,N \tag{7}$$

3.3.3. Sevcik Method

Calculation of fractal dimensions using the Sevcik method (SFD) on N-curves can be written as in Equation (8) [23].

$$SFD = 1 + \frac{\ln(Lc)}{\ln\left[2.(N-1)\right]} \tag{8}$$

Lc in Equation 8 is the total length that has been stated in Equation 6. Another variation of the Sevcik method is the normalization process on the x-axis and y-axis before calculating Lc and SFD. The process of normalization on the x-axis as in Equation 9.

$$x_i^* = \frac{x_i}{x_{\max}} \quad ,i = 1,\ldots,N \tag{9}$$

The value x_i is the value on the initial x-axis, while $x_{\max}$ is the maximum value of x_i. Normalization on the y-axis is expressed as in Equation 10.

$$y_i^* = \frac{y_i - y_{\min}}{y_{\max} - y_{\min}} \quad ,i = 1,\ldots,N \tag{10}$$

y_i is the initial y value, while $y_{\min}$ is the minimum y value while $y_{\max}$ is the maximum value of y.

3.3.4. Variance Method

Variance fractal dimension (VFD) of the signal $s(t)$ is calculated using Hurst Exponent (H) as in Equation 11.

$$H = \lim_{\Delta t\to 0}\left[\frac{\frac{1}{2}\log\left\{\mathrm{var}\left[(\Delta s)_{\Delta t}\right]\right\}}{\log(\Delta t)}\right] \tag{11}$$

H shows the smoothness of the signal. In Equation 11, $(\Delta s)_{\Delta t}$ states $s(t_2)-s(t_1)$ and $\Delta t = t_2 - t_1$. Using Equation 11 VFD can be calculated as in Equation 12.

$$VFD = E + 1 - H \qquad (12)$$

The value of E is the Euclidean dimension, where for the one-dimensional signal the value of E is 1. Thus Equation 12 can be rewritten as Equation 13.

$$VFD = 2 - H \qquad (13)$$

Calculation of VFD can be varied with a value of Δt that varies according to needs. In signal separation with noise, Δt is worth 1 (1 sample of data) while to separate several components of data the value of Δt can be greater.

3.3.5. Higuchi Method

The Higuchi method (HFD) is one of the fractal dimension measurement algorithms that is often used in biomedical signals [26]. The advantages of the Higuchi method are high accuracy and efficiency in measuring fractal dimensions. If a signal with the number of samples N, several lines along the k can be formed, with different resolutions as in Equation 14.

$$X_k^m : x(m), x(m+2k), \dots, x\left(m + \left[\frac{N-m}{k}\right]k\right) \qquad (14)$$

The value of m represents the initial time indication (*m = 1, 2, ..., k*). then, the length of the curve $X_k^m, l_m(k)$ is defined in Equation 15.

$$l_m(k) = \frac{\sum_{i=k}^{\left[\frac{N-m}{k}\right]} \left| x(m+ik) - x(m+(i-1)k) \right| (N-1)}{(N-m/k)k} \qquad (15)$$

Notation $\lfloor a \rfloor$ means floor (a), where $(\lfloor N-m/k \rfloor)k$ is a normalization factor. From Equation 15 the length of the curve for each interval k can be calculated as in Equation 16.

$$L(k) = \sum_{m=1}^{k} l_m(k) \qquad (16)$$

Fractal dimensions are obtained from the slope between plots $\ln(L(k))$ to $\ln(1/k)$. The value is obtained from the relation $L(k) \propto k^{-D}$ where Higuchi fractal (HFD) = D.

3.3.6. Petrosian A and B Method

The Petrosian algorithm calculates the fractal dimension of a signal by changing the signal sequence into a binary sequence. Some variations of this algorithm convert signals into binary sequences in several ways [27] [28]:

a) Petrosian method A = value 1 if x (i)> mean (x), value 0 if x (i) <mean (x)
b) Petrosian Method B = value 1 if x (i)> (mean (x) + std (x)) or x (i) <mean (x) - std (x)
c) Value of 0 if (mean (x) + std (x)> x (i)> mean (x) -std (x))

The petrosian fractal dimension (PFD) is calculated using Equation 17.

$$PFD = \frac{\log_{10}(n)}{\log_{10}(n) + \log_{10}\left(\frac{n}{n + 0.4N_\Delta}\right)} \qquad (17)$$

n is the signal length while $\mathbf{N_\Delta}$ is a sign change in the binary row.

3.3.7. Petrosian C and D Method

The Petrosian algorithm calculates the fractal dimension of a signal by changing the signal sequence into a binary sequence. Some variations of this algorithm convert signals into binary sequences in several ways. In the Petrosian C algorithm, sequential signals are calculated as the difference $\Delta s(t) = s(t+1) - s(t)$. If $\Delta s(t) > 0$ then the value of $\Delta s(t)$ will be set to 1 while $\Delta s(t) < 0$ then the value of $\Delta s(t)$ is set -1. Thus, a binary line of '1' and '-1' will be formed.

In the Petrosian D algorithm, also calculates $\Delta s(t) = s(t+1) - s(t)$. If the value of $\Delta s(t) >$ standard deviation of $s(t)$, the binary series will be '1' and if $\Delta s(t) <$ standard deviation of $s(t)$ then the row will be '-1 '. The petrosian fractal dimension (PFD) is calculated using Equation 18.

$$PFD = \frac{\log_{10}(n)}{\log_{10}(n) + \log_{10}\left(\frac{n}{n + 0.4N_\Delta}\right)} \qquad (18)$$

Where n is the signal length, while $\mathbf{N_\Delta}$ is a sign change in the binary rows.

3.4. Support Vector Machine (SVM)

SVM was developed by Boser, Guyon & Vapnik which was first presented in 1992 at the Annual Workshop on Computational Learning Theory. SVM is one of the alternative learning machines used in solving classification problems using the maximum margin concept [29]. This is what distinguishes the SVM method from other methods of solving classification problems.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SVM is one of the supervised learning methods because the training data set is in the form of input vectors given the target as output. The purpose of this learning is to build a model that can produce the correct output if given new input. The general model used is linear. The linear model is used to separate learning data into two different classes, namely positive and negative classes. This linear model is a decision boundary called a hyperplane.

Suppose $\{x_i, y_i\}$, $i = 1, 2, \ldots, n$ is the set of data pairs as much as n, with $\boldsymbol{x_i} \in \mathbb{R}^m$ and $y_i \in \{+1, -1\}$ is the target. For a real-valued function $f: \mathbb{R}^m \to \mathbb{R}$ then the equation $f(x) = \boldsymbol{w}^T \boldsymbol{x} + b = 0$ is called a hyperplane that separates data into two classes, $\boldsymbol{w} \in \mathbb{R}^m$ is a vector that presents the parameters of weight, and $b \in \mathbb{R}$ is biased. The hyperplane used to classify data is not unique, even in the two-dimensional case.

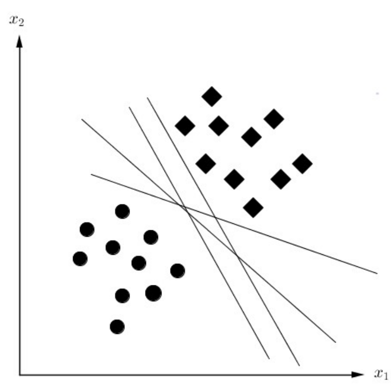


Figure 4: Hyperplane that separates training data

In Figure 4, several hyperplanes are used as training data separators, meaning that there are several choices of parameters w and b, so parameter selection must give the best results. This certainly makes it difficult to choose the best and unique hyperplane. By using SVM, the hyperplane that is used as a decision boundary is a hyperplane that groups data with maximum margins. Thus, to get the optimal hyperplane $\boldsymbol{f(x)} = \boldsymbol{w^T x} + \boldsymbol{b} = \boldsymbol{0}$ that is a hyperplane that can classify data correctly, then the values w and b must be searched through the margin optimization process. Figure shows an illustration of a hyperplane that maximizes margins

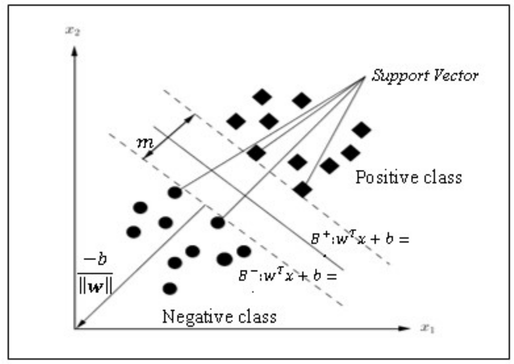


Figure 5: Hyperplane with Maximum Margin

In the application, not all data can be separated linearly by the hyperplane even though it has used soft margins. The soft margin allows some data to be on the wrong side of the decision boundary or to provide softness for some data that is misclassified. To realize this soft margin, the slack variable is introduced. However, not all data can be separated linearly by the hyperplane even though it has used soft margin by adding a slack variable. To overcome this, the Kernel method is used by mapping data to higher dimensions so that the data can be separated linearly or it is also linearly separable [30].

Several kernel functions are often used in SVM literature, among others [17]:

1) **Linear Kernel**, is the simplest kernel function which is the dot product of two vectors. Linear Kernel functions are defined in Equation 19.
$$K(x, x') = x^T x' \quad (19)$$
2) **Polynomial Kernel**, with a degree p, where p is the natural number that can be defined in the Equation 20.
$$K(\boldsymbol{x}, \boldsymbol{x'}) = (\boldsymbol{x^T x'} + 1)^p \quad (20)$$
3) **Radial Base Gaussian Kernel**, also called Gaussian Kernel function.
4) **Tangent Hyperbolic Kernel**, is a kernel that is often used for neuron networks.

In classification, the use of the kernel method is used to obtain optimal functions as a decision boundary in the form of soft margins. Explicitly, the advantage of using the kernel is that it does not require treatment in high-dimensional space. This technique is referred to as the kernel trick. Selection of kernel functions will also affect the results of accuracy in classification.

3.5. K-Fold Cross Validation

In the SVM classification, after obtaining the best SVM classifier function model, the model

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

evaluation will be carried out. K - Fold Cross Validation is used to calculate the accuracy of the classification function model for new data. To present the results of K-Cross Validation, a confusion matrix is used. Confusion matrix contains the number of elements that have been correctly or incorrectly classified for each class [31]. A confusion matrix is usually used to calculate accuracy in the concept of data mining and is used to present the results of K-Fold Cross Validation with the contents of the number of true positives (TP), true negative (TN), false positive (FP), and false-negative (FN) as in Equation 21 [31].

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \times 100\% \qquad (21)$$

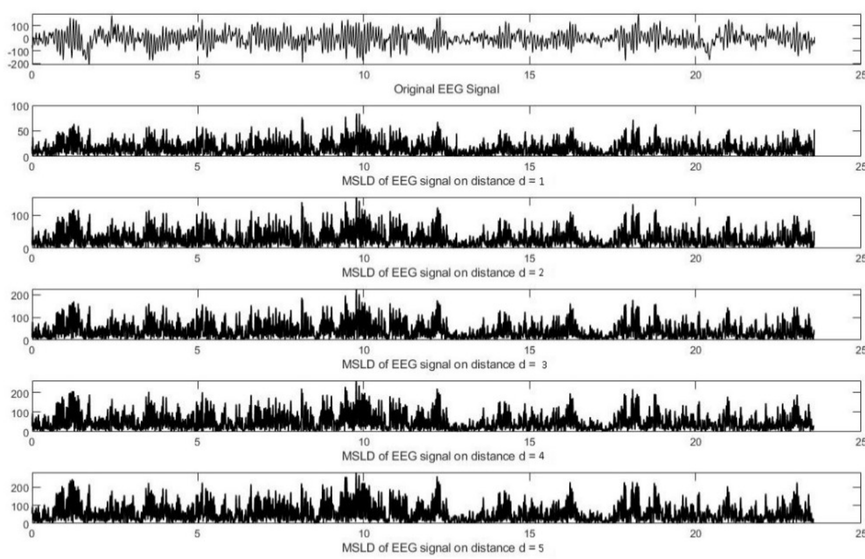
4. RESULTS AND DISCUSSION

In this research, EEG signal data were used from the Bonn University dataset which had 3 classes namely normal (set O), epileptic no seizure (set N), and seizure conditions (set S). This fractal analysis of EEG signals is expected to be able to distinguish EEG signals in seizure, no seizure, and normal conditions. The multiscale method used is a multi-distance signal level difference (MSLD). Signal analysis using MSLD is combined with the fractal dimension. Defining signal complexity with chaotic approaches is to use fractal dimensions. The fractal dimensions used in this research are Box counting (BC), Higuchi (HFD), Katz, Petrosian C (PetC), Petrosian D (Pet D), Sevcik (SVD) and Variance (VFD) methods. After the fractal dimension is done, then classification use SVM.

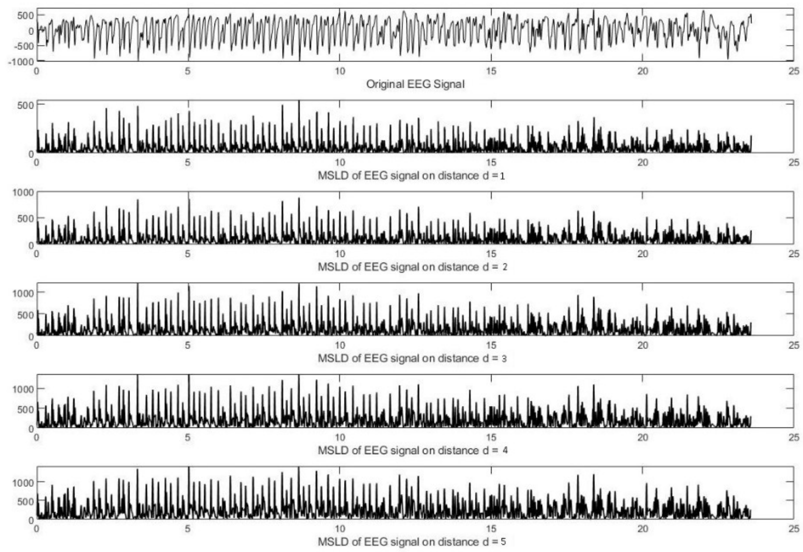
4.1. Analysis of Multi-distance Signal Level Difference (MSLD)

The results of the MSLD process for EEG signals with distance d = 1 - 5 in the 3 classes used can be shown in Figure 6. MSLD calculated the difference in absolute values of 2 samples of data at a distance d. So, the results of the signal are always in the form of positive values.

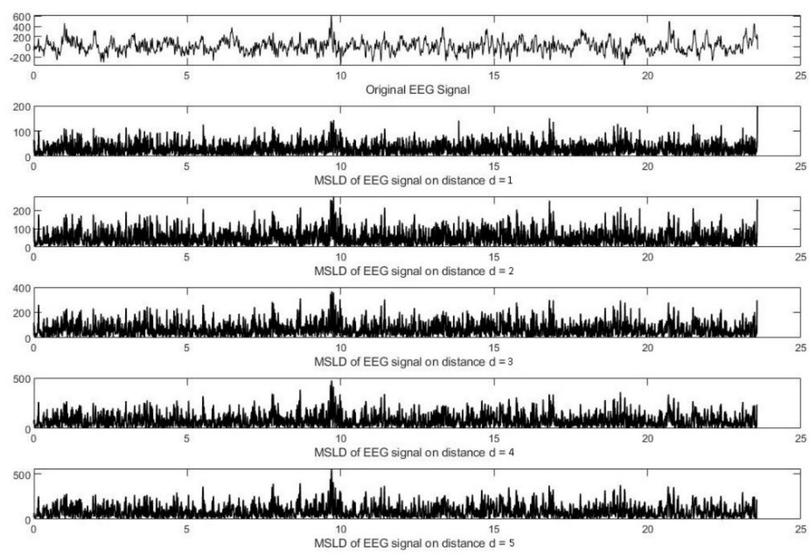
In Figure 6, it can be seen that the signal generated is always positive because of the absolute sign in Equation (2). In Figure 6 (c), it can be seen that the Seizure signal has a high enough signal amplitude because of the seizure in that condition. In Figure 6 (b), the interictal EEG signal condition is different from the normal EEG signal. Although they both have a lower amplitude than the seizure, in general the ranges of the two are different. The MSLD process is expected to strengthen differences between classes and strengthen inter-class similarities.



(a)



(b)



(c)

Figure 6: MSLD EEG Signals (a) Normal, (b) No Seizure, (c) Seizure

4.2. Analysis of Fractal Dimension

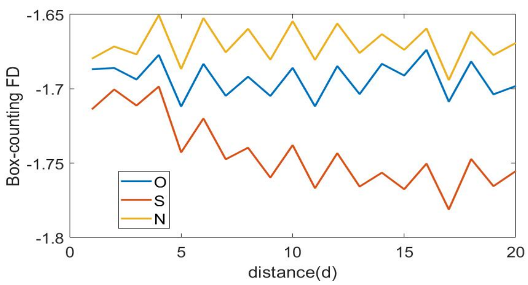
Furthermore, the results of the fractal analysis on EEG signals using MSLD are combined with the fractal dimension. The fractal dimension will define signal complexity with chaotic approaches. The fractal dimensions used in this research are Box counting (BC) method, Higuchi method (HFD), Katz method, Petrosian C (PetC), Petrosian D (Pet D), Sevcik method (SVD), and variance method (VFD). The results of the fractal dimension are shown in Figure 7. Based on the visual results of the fractal dimension, the EEG seizure signal produces the highest fractal value and the epileptic no seizure EEG signal is the lowest. Each fractal dimension method used produces significantly different results so that the classification process can be determined properly.

Analysis of the fractal dimensions of MSLD is carried out at the distance d = 1-20, then there will be 20 fractal dimensions produced as a

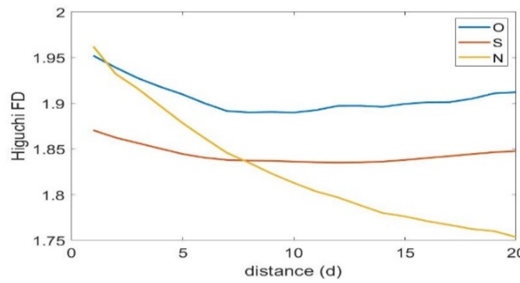
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

feature of each fractal dimension. The results of each fractal dimension at a distance of d = 1 - 20 are shown in Figure 7. In Figure 7 (a), (f) and (g) produce different features for the Seizure (S) and are separate from the other two classes. Meanwhile, in Figures 7 (b), (c), (d) and (e) produce features that are slightly coincided or crossed so that there is the possibility of misclassification.

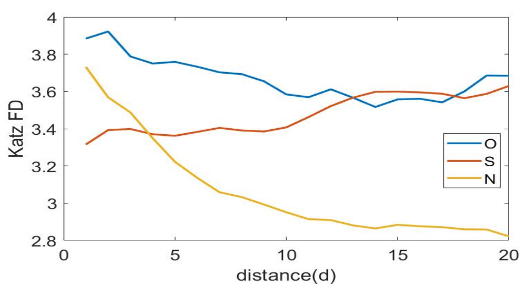
Visually, SVD and VFD produce the best characteristics because the three classes are strictly separate. The average values of the fractal dimensions are shown so they appear separate. However, there are still possible errors in classification. Thus, the classification is carried out using SVM with various kernels.



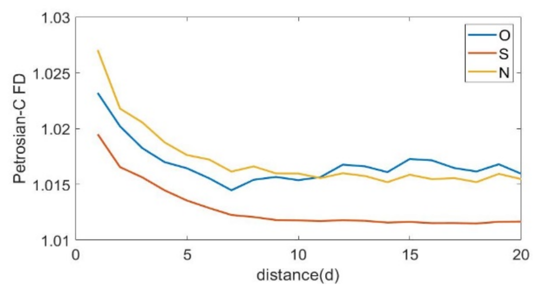
(a) Box Counting



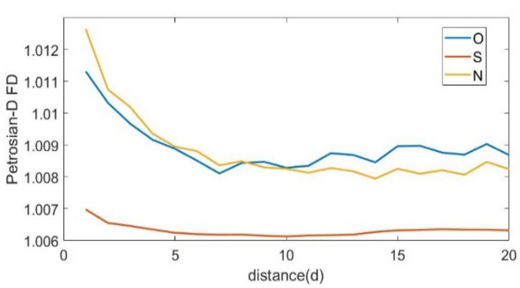
(b) Higuchi Method



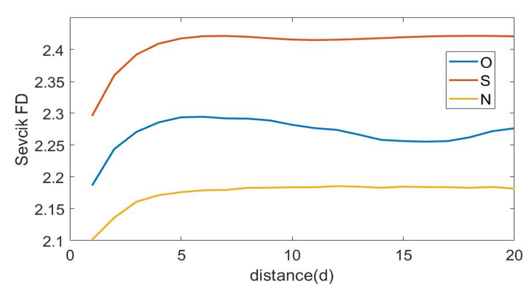
(c) Katz Method



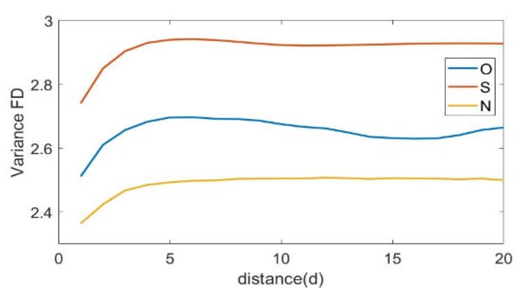
(d) Petrosian C



(e) Petrosian D



(f) Sevcik Method



(g) Variance Method

Figure 7: Fractal Dimensions on EEG Signals

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.3. Classification Analysis using Support Vector Machine

The next process is the performance test of the MSLD fractal dimension using SVM with Linear, Quadratic, and Gaussian kernels. The results of the classification accuracy using SVM are shown

Table 1. The highest accuracy was achieved by MSLD with the fractal dimension of the Variance

in The highest accuracy was achieved by MSLD with the fractal dimension of the Variance method using quadratic SVM at 90%. It appears that reducing the number of distances used in MSLD can improve accuracy.

method using quadratic SVM at 90%. It appears that reducing the number of distances used in MSLD can improve accuracy.

Table 1: Accuracy (%) Using SVM and 5-fold CV

	Linear SVM			Quadratic SVM			Medium Gaussian SVM		
	Scale 1-20	Scale 1-15	Scal e 1-10	Scale 1-20	Scale 1-15	Scale 1-10	Scale 1-20	Scale 1-15	Scale 1-10
Box Counting	70.7	72.3	68.7	66.7	70.3	72.7	70.7	72.3	73
Higuchi Method	92.3	93.3	92.7	93.7	95	95.7	91.7	92.3	91.3
Katz Method	87.3	86.7	81.3	84.3	84.3	81.7	88.3	85.7	79
Petrosian C	95.7	96	95.3	95	96.7	97	96.7	95.3	94.7
Petrosian D	94	94	92.7	94.7	94.7	96	94.3	94.3	93
Sevcik Method	98.3	98	97.7	98	98	98.3	98	98	96.7
Variance Method	98	98.7	98.7	98.7	98.7	99	98.3	98.3	94.7

MSLD shows general occurrences of two samples over a certain distance range. This result shows some differences in features between classes. The advantage of MSLD is that the signal variance value does not change, different from the coarse-grained procedure which reduces the signal variance value as discussed in previous studies [14]. The loss of MSLD is the distance range to be calculated is determined by trial and error. However, empirically, MSLD is good for distance ranges d = 1-15.

The results of this research have high accuracy compared to previous research using MSLD with SampEn. The results of previous research using MSLD with SampEn had a high accuracy of 97.7% [15]. This shows that MSLD by applying various fractal dimensions can provide higher accuracy results in the classification of EEG signals. The MSLD fractal method which has the highest accuracy is the variance method compared to other dimensional fractal methods. Fractal dimension using variation method with classification using SVM Quadratic has higher accuracy than other fractal dimension methods, namely on a scale of 1-20 and 1-15 of 98.7% and 1-10 of 99%. This happens, because the calculation of fractal dimensions using the variance method produces the exact characteristics of the EEG signal as input for the classification process. The MSLD method can be further developed in combination with various other feature extraction methods such as entropy calculations, statistics, or signal complexity methods.

5. CONCLUSIONS

In this research, the classification of epileptic EEG signals was done by using fractal dimension MSLD. There are three classifications of EEG signals namely normal, no seizure, and seizure. Several fractal dimension methods used produce differences, but in general, the EEG seizure signal produces the highest fractal value and the epileptic no seizure EEG signal is the lowest. Furthermore, the classification process carried out using Quadratic SVM and 5-fold CV resulted in an accuracy of 98.7% with a scale of 1-20 and scale 1-15, and an accuracy of 99% on a scale of 1-10. So, it can be concluded that using MSLD fractal dimensions with Quadratic SVM produces the highest accuracy of 99% on a scale of 1-10. For further research, it is recommended that MSLD be used to classify other biopotential signals that have high complexity

6. ACKNOWLEDGEMENTS

The authors would like to thank all the members of the research team for their collaboration. Thank you to Telkom University for providing our support to be able to conduct research in accordance with scientific fields.

REFRENCES:

[1] J. Kevric and A. Subasi, "Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system," *Biomed. Signal Process. Control*, 2017, doi: 10.1016/j.bspc.2016.09.007.

[2] D. E. Hernández, L. Trujillo, E. Z-Flores, O. M. Villanueva, and O. Romo-Fewell, "Detecting Epilepsy in EEG Signals Using Time, Frequency and Time-Frequency Domain Features," in *Studies in Systems, Decision and Control*, 2018.

[3] E. C. Ifeachor, G. T. Henderson, C. Goh, H. S. K. Wimalaratna, and N. Hudson, "Biopattern Analysis and Subject-Specific Diagnosis and Care of Dementia," 2006, doi: 10.1109/iembs.2005.1616974.

[4] M. N. Tibdewal, H. R. Dey, M. Mahadevappa, A. K. Ray, and M. Malokar, "Multiple entropies performance measure for detection and localization of multi-channel epileptic EEG," *Biomed. Signal Process. Control*, 2017, doi: 10.1016/j.bspc.2017.05.002.

[5] P. F. Lin, M. T. Lo, J. Tsao, Y. C. Chang, C. Lin, and Y. L. Ho, "Correlations between the signal complexity of cerebral and cardiac electrical activity: A multiscale entropy analysis," *PLoS One*, 2014, doi: 10.1371/journal.pone.0087798.

[6] A. Mirzaei, A. Ayatollahi, P. Gifani, and L. Salehi, "EEG analysis based on wavelet-spectral entropy for epileptic seizures detection," 2010, doi: 10.1109/BMEI.2010.5639894.

[7] I. Wijayanto, A. Rizal, and S. Hadiyoso, "Multilevel Wavelet Packet Entropy and Support Vector Machine for Epileptic EEG classification," 2018, doi: 10.1109/ICSTC.2018.8528634.

[8] A. C. Yang *et al.*, "Cognitive and neuropsychiatric correlates of EEG dynamic complexity in patients with Alzheimer's disease," *Prog. Neuro-Psychopharmacology Biol. Psychiatry*, 2013, doi: 10.1016/j.pnpbp.2013.07.022.

[9] C. Goh, B. Hamadicharef, G. Henderson, and E. Ifeachor, "Comparison of Fractal Dimension Algorithms for the computation of EEG Biomarkers for Dementia," in *2nd International Conference on Computational Intelligence in Medicine and Health Care (CIMED2005)*, 2005, pp. 464–471.

[10] M. Phothisonothai and K. Watanabe, "Optimal Fractal Feature and Neural Network: EEG Based BCI Applications," in *Brain-Computer Interface Systems - Recent Progress and Future Prospects*, 2013.

[11] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series.," *Phys. Rev. Lett.*, vol. 89, no. 6, p. 068102, 2002, doi: 10.1103/PhysRevLett.92.089803.

[12] A. Rizal, R. Hidayat, and H. A. Nugroho, "Comparison of Multiscale Entropy Techniques for Lung Sound Classification," *Indones. J. Electr. Eng. Comput. Sci.*, 2018, doi: 10.11591/ijeecs.v12.i3.pp984-994.

[13] A. Humeau, G. Mahé, F. Chapeau-Blondeau, D. Rousseau, and P. Abraham, "Multiscale analysis of microvascular blood flow: A multiscale entropy study of laser doppler flowmetry time series," *IEEE Trans. Biomed. Eng.*, 2011, doi: 10.1109/TBME.2011.2160865.

[14] A. Rizal, R. Hidayat, and H. A. A. Nugroho, "Hjorth Descriptor Measurement on Multidistance Signal Level Difference for Lung Sound Classification," *J. Telecommun. Electron. Comput. Eng.*, 2017.

[15] A. Rizal and S. Hadiyoso, "Sample Entropy on Multidistance Signal Level Difference for Epileptic EEG Classification," *Sci. World J.*, 2018, doi: 10.1155/2018/8463256.

[16] A. Rizal, R. Hidayat, and H. A. Nugroho, "Fractal Dimension for Lung Sound Classification in Multiscale Scheme," *J.*

Comput. Sci., 2018, doi: 10.3844/jcssp.2018.1081.1096.

[17] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag New York Inc., 2006.

[18] A. Shoeb and J. Guttag, "Application of machine learning to epileptic seizure detection," in *In: Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 975–82.

[19] Ubeyli, I. Guler, and E. Derya, "Multiclass support vector machines for EEG-signals classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 2, pp. 117–126, 2007.

[20] 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 - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.*, 2001, doi: 10.1103/PhysRevE.64.061907.

[21] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A Comparative Study of Texture Measures for Terrain Classification," *Syst. Man Cybern. IEEE Trans.*, vol. SMC-6, no. 2, pp. 269–285, 1976.

[22] R. P. Burn and B. B. Mandelbrot, "The Fractal Geometry of Nature," *Math. Gaz.*, 2007, doi: 10.2307/3615422.

[23] C. Sevcik, "A procedure to estimate the fractal dimension of waveforms," *Complex. Int.*, 1998.

[24] B. B. Mandelbrot, *The Fractal Geometry of Nature*. New York: W. H. Freeman and Company, 1983.

[25] M. J. Katz, "Fractals and the analysis of waveforms," *Comput. Biol. Med.*, 1988, doi: 10.1016/0010-4825(88)90041-8.

[26] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Phys. D Nonlinear Phenom.*, 1988, doi: 10.1016/0167-2789(88)90081-4.

[27] A. Petrosian, "Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns," 2002, doi: 10.1109/cbms.1995.465426.

[28] R. Esteller, G. Vachtsevanos, J. Echauz, and B. Litt, "A Comparison of waveform fractal dimension algorithms," *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, 2001, doi: 10.1109/81.904882.

[29] I. Guyon, V. Vapnik, B. Boser, L. Bottou, and S. A. Solla, "Capacity control in linear classifiers for pattern recognition," 1992, doi: 10.1109/ICPR.1992.201798.

[30] A. Scholkopf, B & Smola, *Learning with Kernels – Support Vector Machines, Regularization, Optimizationand Beyond*. Cambridge: The MIT Press, 2002.

[31] F. Gorunescu, "Data mining: Concepts, models and techniques," *Intell. Syst. Ref. Libr.*, 2011, doi: 10.1007/978-3-642-19721-5.