Advanced Science Letters
Vol. 23, 3972–3974, 2017

Electrocardiogram Signal Classification Using Higher-Order Complexity of Hjorth Descriptor

Sugondo Hadiyoso[1,*] and Achmad Rizal[2]

[1]*Telkom Applied Science School, Telkom University, Bandung, Indonesia*
[2]*School of Electrical Engineering, Telkom University, Bandung, Indonesia, Telkom University, Bandung 40257, Indonesia*

ECG signal is a bio-potential signal generated by the heart muscle that can be used to detect heart abnormalities. Research on the ECG signal classification becomes a topic which is done mostly by researchers. The goal is to find the simplest algorithm, less computation but still has a good performance. In this research, the Higher Order Complexity of Hjorth Descriptor is used to extract the feature of ECG signal. The testing data consists of three types of ECG signal namely Normal Sinus Rhythm (NSR), Atrial Fibrillation (AF) and Congestive Heart Failure (CHF). K-Nearest Neighbor (K-NN) and multilayer perceptron (MLP) is used for classification the feature of the signal from Hjorth Descriptor result. Our propose method produce 94% accuracy using both MLP and K-NN respectively.

Keywords: ECG, Hjorth Descriptor, Complexity, K-NN, MLP.

1. INTRODUCTION

Electrocardiogram (ECG) is a biological signal arising from the electrical activity of the heart. The electrical signals of heart trigger the onset of heart muscle contraction that pumps blood throughout the body.[1] The ECG signal representing heart health. Abnormalities of the heart can be detected by looking at the shape, rhythm, and the orientation of the ECG signal.[1] ECG signal has a specific shape so that by looking at the change in shape of the ECG signal can be seen the part of heart's problem.[2]

Many digital signal processing techniques developed to be able to recognize, determine and classify the pattern of the ECG signal automatically. Digital signal processing techniques in the time domain become the primary choice because it does not require transformation to another domain. One of the methods in the time domain is often used in lung sounds analysis is Principal component analysis (PCA).[3–5]

Another method that often used is Hjorth descriptor.[6] Hjorth descriptor was originally used for the analysis of EEG signals and then used for EMG analysis,[7] ECG[8] and the sound of breathing.[9] The Hjorth descriptor consists of three parameters: activity, mobility and complexity which considered representing the characteristics of a signal. In previous research,[8] we used first order complexity of Hjorth descriptors. The use of higher-order complexity has not been shown in previous studies. In this work, N-order complexity is used for ECG signal feature extraction. N-Order Complexity is calculated from the derived signals from 1 to $N+1$.

By using N-order complexity order, the difference between the classes of the ECG signal is expected to be increasingly evident thus producing higher accuracy in ECG signal classification.

2. MATERIAL AND METHOD

2.1. Data of ECG Signal

ECG signal data used was taken from the Physionet database.[10] The same data used in our previous study using the descriptor Hjorth.[8] The data used consists of three classes of Normal Sinus Rhythm (NSR), Atrial fibrillation (AF) and Congestive Heart Failure (CHF). Each class contains 50 data with length 2–3 QRS and sampling frequency of 250 Hz. NSR is a normal condition in which the ECG signal is formed to have a pattern and a rhythm.[11] In AF signal occurs due to the chaotic rhythm uncoordinated atrial activation. The condition is caused by degeneration of atril.[11] CHF is a condition in which the heart loses its pumping ability of the heart effectively. The heart pumps less blood to pump so it does not meet the needs of the whole body.[12] In each data, we did two processes, the data were normalized the

*Author to whom correspondence should be addressed.

1936-6612/2017/23/3972/003 doi:10.1166/asl.2017.8251

amplitude and remove the DC component. The first process is done with the equation $x_{dc} = x(n) - \text{mean}(x)$. The next process is the amplitude normalization, $x(n) = x_{dc}(n)/\max|x_{dc}|$ which makes the value of $x_n(n)$ lies between -1 and $+1$.

2.2. Hjorth Descriptor

Hjorth descriptor is a parameter to quantize EEG characteristic.[6] Hjorth descriptor consists of activity, mobility, and complexity.[6] If we have $x(n)$, the input signal, then σ_0 = standard deviation of $x(n)$. For $x_1(n) = x(n) - x(n-1)$ we will have σ_1 = variance of $x_1(n)$. Meanwhile, σ_2 = variance of $x_2(n)$, where $x_2(n) = x_1(n) - x_1(n-1)$ or generally can be formulated as:

$$x_N(n) = x_{n-1}(n) - x_{n-1}(n-1) \tag{1}$$

The Hjorth descriptor then expressed as in Eqs. (2)–(4).[13]

$$\text{Activity} = \sigma_0^2 \tag{2}$$

$$\text{Mobility} = M = \sigma_1^2/\sigma_0^2 \tag{3}$$

$$\text{Complexity order of } n = \sqrt[2]{\left(\frac{\sigma_{n+1}^2}{\sigma_n^2} - \frac{\sigma_n^2}{\sigma_{n-1}^2}\right)} \tag{4}$$

In this study only complexity parameter that is used as a feature. Complexity order 1 to order 5 was used as features for input of classifier. Order 1 to order 5 are selected as characteristics due to the order 6 and higher produce complexity value close to zero or tends to be flat for all data.

3. RESULTS AND DISCUSSION

Figure 1 shows $x_N(n)$ with $N = 1$ to 5 for the CHF signal. It is seen that signals for N-derivative have relatively flat amplitude but more dynamic. The same pattern is shown in Figure 2. From both of them, it can be seen that for N-derivatives makes the differences of the signal will be more evident.

To test the performance the results of feature extraction, we use MLP and K-NN as classifier. Since both of them are supervised learning classifier, we divide training data and testing data using 10-fold cross-validation, five-fold cross-validation and 50%–50% data separation. Figure 3 displays the effect of the number of

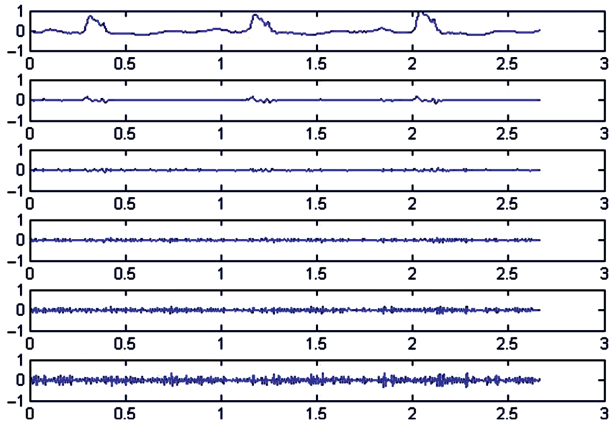

Fig. 1. CHF Signal and its N-derivative for $N = 1, 2, \ldots, 5$.

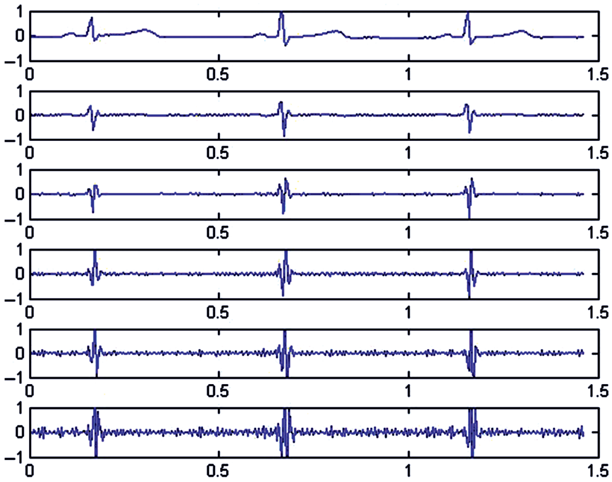

Fig. 2. NSR Signal and its N-derivative for $N = 1, 2, \ldots, 5$.

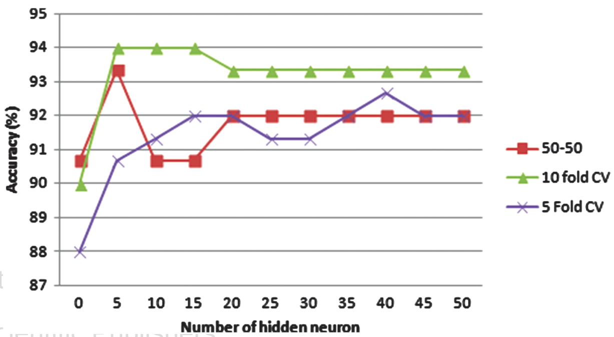

Fig. 3. Effect of the number of hidden neurons in MLP to accuracy (10 fold CV).

hidden neurons in MLP as to the accuracy of each data division scheme. The highest accuracy is achieved using 10-fold CV obtained 94%.

Some parameters used for performance measurement are accuracy, sensitivity, and specificity. All the parameters are expressed as:

$$\text{Sensitivity (SE)} = TP/(TP+FN) \tag{5}$$

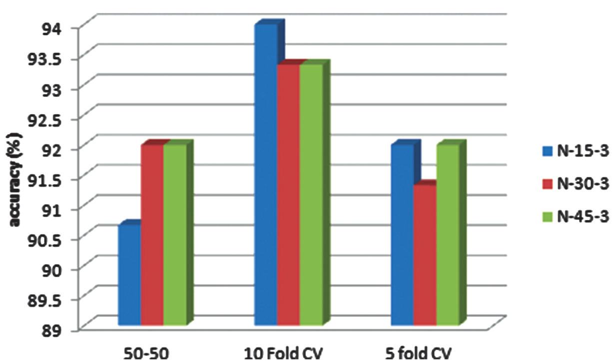

Fig. 4. Accuracy using MLP with some configuration and distribution of training data and test data (the highest result is 94%).

Adv. Sci. Lett. 23, 3972–3974, 2017

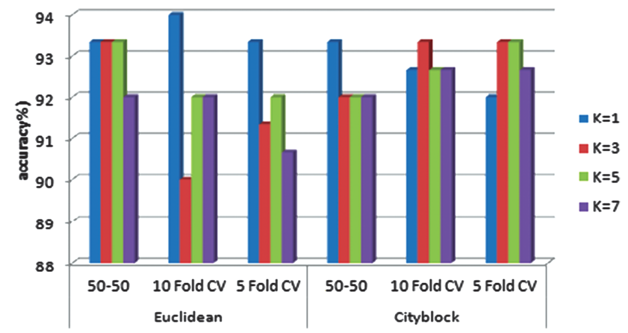

Fig. 5. Accuracy using K-NN with some configuration and distribution of training data and test data (the highest result is 94%).

$$\text{Specificity (SP)} = TN/(TN+FP) \quad (6)$$

$$\text{Accuracy (ACC)} = (TP+TN)/(TP+FN+TN+FP) \quad (7)$$

Where TP: true positive, TN: true negative, FN: false negative, and FP: false positive.

Figure 4 shows more evident comparison results in accuracy between the three configurations of MLP. The best results were achieved by MLP with the configuration N-15-3 and using 10-fold CV.[14] Meanwhile, the same result is shown in Figure 5. The highest result was also obtained for a 10-fold CV on the K-NN with $K = 1$ and Euclidean distance method.

Table I shows the classification error that occurred. NSR signals generate 100 % SP and SE, which means that none of the NSR data misidentified and no data from another class known as NSR. Meanwhile, the most common mistake is CHF signals that are classified as AF. in AF; the abnormalities frequently appear are rhythm and QRS complex that often not arise, while the CHF the main difference is usually in the QRS complex.[2]

Compared with previous work that uses Hjorth descriptors, the results are slightly worse.[8] By using activity, mobility and complexity order 1 in an earlier research can produce up to 100% accuracy. Activity, mobility and first order complexity describe variance of the signal, first-order variation of signal and second-order variation of the signal. Meanwhile in this research we only use second-order and higher-order variation of ECG signal. It is mean that the signal tent to be flatter than the original signal. Even in this research, the results are slightly descend, but technically, the algorithm can produce sufficient the number of characteristics with good consistency. The use of higher-order complexity of Hjorth descriptor will generate features that are multi-scale.[15] Multi-scale nature is one of the fundamental properties of biological signals that a complex system.[15] With the results obtained provide the possibility of applying the feature extraction on other biological signals using Higher-order complexity of Hjorth descriptor.

Table I. Confusion matrix, SP, SE and accuracy at the best accuracy.

	Classified as					
Data	NSR	AF	CHF	Se (%)	Sp (%)	Acc (%)
NSR	50	0	0	100	100	94%
AF	0	48	2	96	93	
CHF	0	7	43	86	98	

Note: Se = sensitivity, SP = specificity, Acc = accuracy.

4. CONCLUSIONS

On this research has presented the use of Higher-order Complexity of Hjorth descriptor for feature extraction to the ECG signals classification. In the three data classes (NSR, AF and CHF), the maximum accuracy achieved 94%. This result is slightly lower when compared with the use of three parameters of Hjorth descriptors as in previous work. By using Higher-order Complexity of Hjorth descriptors is able to show the dynamics of the signal on the N-order signal derivative. In this study, the data classes of ECG signals that are used tend common abnormalities. The next challenge is how the proposed method can detect abnormal ECG signals which are rare or uncommon.

Acknowledgments: This work was supported in part by Kemristekdikti under Hibah Bersaing research grant.

References and Notes

1. W. J. Tompkins, Electrocardiography, in Biomedical Digital Signal Processing, edited by W. J. Tompkins, Prentice Hall, New Jersey **(2000)**, pp. 24–54.
2. T. A. M. Brosche, The EKG Handbook, Jones and Bartlett Publisher **(2010)**.
3. A. Bollmann, M. Roig, F. Castells, P. Laguna, and S. Leif, *EURASIP J. Adv. Signal Process* 2007 **(2007)**.
4. Q. U. Xiao, C. Wei, and G. D. Fei, ECG signal classification based on BPNN, *International Conference on Electrical Information and Control Engineering (ICEICE)* **(2011)**, Vol. 2011, pp. 1362–1364.
5. R. Joy, U. R. Acharya, and L. Choo, *Biomed. Signal Process. Control* 8, 437 **(2013)**.
6. B. Hjorth, *Clin. Neurophysiol.* 29, 306 **(1970)**.
7. M. Mouzé-Amady and F. Horwat, *Electroencephalogr. Clin. Neurophysiol.—Electromyogr. Mot. Control* 101, 181 **(1996)**.
8. Rizal and S. Hadiyoso, ECG signal classification using Hjorth descriptor, *2015 International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT)* **(2015)**, pp. 3–6.
9. A. Rizal, R. Hidayat, and H. A. Nugroho, Determining lung sound characterization using Hjorth descriptor, *2015 International Conference on Control, Electronics, Renewable Energy and Communication (ICCEREC)* **(2015)**, pp. 54–57.
10. Physionet.org, ECG Database, [Online], Available: http://physionet.org/physiobank/database/#ecg.
11. R. Alcaraz and J. J. Rieta, *Biomed. Signal Process. Control* 5, 1 **(2010)**.
12. J. L. Schuster, C. Spence, M. Jacobs, and A. Wilkonson, Living with advanced congestive heart failure: A guide for family caregivers, The Washington Home Center for Palliative Care Studies **(2002)**.
13. B. Hjorth, *Electroenchepalography Clin. Neurophysiol.* 34, 321 **(1973)**.
14. F. Albu, A. Mateescu, and N. Dumitriu, Architecture selection for a multilayer feedforward network, *International Conference on Microelectronics and Computer Science*, Chishinau, Republic of Moldavia **(1997)**, pp. 131–134.
15. M. Costa, A. L. Goldberger, and C.-K. Peng, *Phys. Rev. Lett.* 89, 068102 **(2002)**.

Received: 26 February 2016. Accepted: 19 April 2016.