



Klasifikasi Sinyal Elektrokardiogram Menggunakan Renyi Entropy

Nano Estananto¹, Achmad Rizal²

¹Telkom University, email: estananto@telkomuniversity.ac.id

²Telkom University, email: achmadrizal@telkomuniversity.ac.id

Abstrak

Sinyal Elektrokardiogram (EKG) adalah salah satu indikator penting tentang aktivitas kelistrikan jantung. Sinyal ini dapat menjadi basis analisa kerja jantung sehingga menjadi penting. Sampai saat ini sinyal EKG tidak dapat digambarkan dengan persamaan matematis yang tepat karena sinyal EKG berulang tapi tidak tepat periodik. Untuk itu digunakan signal complexity untuk menjelaskan perilaku sinyal EKG. Renyi entropy adalah salah satu matrik untuk menganalisa kompleksitas sinyal, dan merupakan bentuk umum dari entropy lain. Dengan demikian akurasi yang dihasilkan lebih tinggi. Pada penelitian ini menggunakan Renyi Entropy dihasilkan akurasi 100% untuk tiga kelas data EKG. Digunakan Support Vector Machine (SVM) linear, qubic, dan quadratic sebagai pembanding satu sama lain. Keunggulan metode ini adalah akurasi yang tinggi dengan menggunakan jumlah ciri yang jauh lebih sedikit.

Kata kunci: EKG, SVM, Renyi Entropy

Abstract

Electrocardiogram (ECG) signals are one of the important indicators of heart electrical activity. This signal can be the basis for analyzing the work of the heart so that it becomes important. Until now the ECG signal cannot be described with the right mathematical equation because the ECG signal is repeated but not periodically correct. For this reason, complexity signals are used to explain the behavior of ECG signals. Renyi entropy is a matrix for analyzing signal complexity, and is a common form of other entropy. Thus the resulting accuracy is higher. In this study using Renyi Entropy produced 100% accuracy for three classes of ECG data. Linear, Qubic, and Quadratic Support Vector Machine (SVM) are used to be compared to each other. The advantage of this method is high accuracy by using a number of traits that are far less.

Keywords: ECG, SVM, Renyi Entropy

1. Pendahuluan

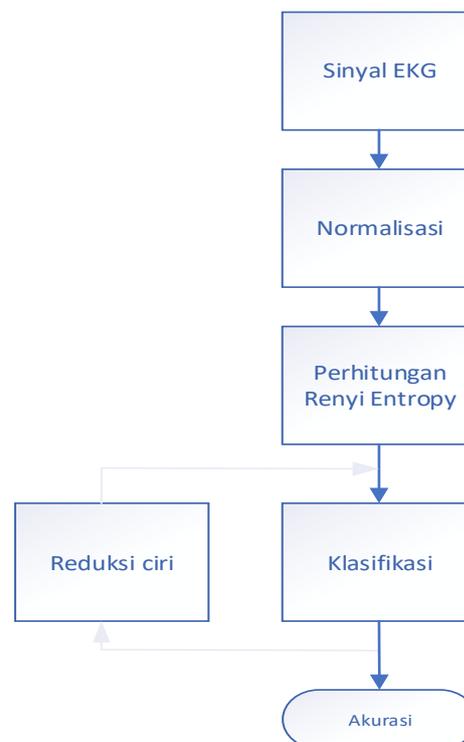
Elektrokardiogram (EKG) adalah sinyal biologi yang dihasilkan dari aktifitas kelistrikan jantung . Sinyal ini timbul untuk memicu kontraksi otot jantung yang akan memompa darah ke seluruh tubuh [1]. Sinyal EKG merupakan representasi tingkat kesehatan jantung seseorang

yang digambarkan lewat irama, bentuk dan orientasi sinyal EKG [2]. Untuk mengurangi subjektivitas dalam menilai pola sinyal EKG, berbagai metode pengoahan sinyal digital dikembangkan oleh para peneliti berdasarkan karakteristik sinyal EKG. Salah satu sifat sinyal biologi seperti sinyal EKG adalah signal complexity, yaitu istilah yang mengacu pada kesulitan untuk memprediksi perilaku sinyal atau kesulitan untuk menjelaskan sifat sinyal secara keseluruhan [3]. Signal complexity muncul sebagai bentuk adaptasi system terhadap perubahan yang terjadi [3]. Sebagai contoh sinyal EKG berulang tetapi tidak tepat periodic sehingga dinamakan dengan quasi-periodic. Sampai saat ini tidak ada satupun persamaan matematis yang dapat menggambarkan sinyal EKG secara tepat. Sebagai gantinya digunakan beberapa metric yang menggambarkan signala complexity tersebut.

Salah satu metric untuk signal complexity adalah entropy. Pada mulanya Shannon menggunakan entropy untuk menjelaskan kandungan informasi pada suatu sinyal komunikasi [4]. Seiring perkembangan teknologi, bermunculan metode entropy yang lain seperti Reny entropy [5], Tsallis entropy [6], wavelet entropy [7] dan lain-lain. Entropy juga digunakan mengkuantisasi perilaku sinyal biologi seperti pada EEG [8], suara paru [9], dan EKG [10].

Pada penelitian ini digunakan Renyi entropy pada banyak orde untuk ekstraksi ciri pada klasifikasi sinyal EKG. Pada penelitian sebelumnya multiscale Renyi entropy digunakan untuk mengenali heart rate variability (HRV) [10a]. Hasil yang didapat menunjukkan bahwa Renyi entropy lebih baik dalam membedakan pola RR pada sinyal EKG normal, dan EKG pada Cardiac Autonomic Neuropathy (CAN). Pada penelitian ini Renyi entropy digunakan untuk klasifikasi tiga jenis sinyal EKG: normal, atrial fibrillation (AF), dan congestive heart failure (CHF). Renyi entropy dihitung pada orde $q = 2 - 100$ sebagai ciri tiap data dan ditambah dengan Shannon entropy dari sinyal. Sebagai classifier digunakan support vector machine dengan beberapa variasi. Tujuan yang ingin dicapai adalah jumlah orde atau ciri paling sedikit yang bisa menghasilkan akurasi klasifikasi sinyal EKG tertinggi. Diharapkan hasil yang didapat bisa digunakan sebagai rekomendasi metode ekstraksi ciri untuk sinyal EKG.

2. Bahan dan Metode



Gambar 1. Diagram alir metode yang digunakan

Proses yang digunakan dalam penelitian ini seperti ditunjukkan oleh Gambar 1. Mula-mula sinyal EKG dinormalisasi untuk emnyamakan level tegangan dan menghilangkan komponen DC. Selanjutnya dilakukan pengukuran Renyi entropy sebagai ciri. Proses selanjutnya adalah klasifikasi menggunakan support vector machine (SVM) dengan beberapa kernel untuk mendapatkan akurasi. Jumlah ciri yang digunakan kemudian dikurangi untuk menguji pengaruh jumlah ciri terhadap akurasi. Penjelasan lebih detail tiap proses dipaparkan pada bagian berikut ini.

2.1 Data Sinyal EKG

Data sinyal EKG diambil dari MIT-BIH database yang disimpan di physionet.org [11]. Sinyal EKG yang digunakan terdiri dari tiga kelas yaitu normal sinus rhythm (NSR), Congestive heart failure (CHF), dan Atrial fibrillation (AF) yang masing masing berjumlah 50 data dengan frekuensi sampling 250 Hz. Tiap data mempunyai panjang 2-3 detik sehingga tidak data memiliki satu sampai dua QRS. NSR mewakili sinyal EKG normal sedangkan AF dan CHF mewakili sinyal EKG abnormal [12][13].

Pada sinyal EKG dilakukan proses normalisasi sebagai berikut. Jika sinyal EKG $x(n)$ dengan $n = 1, 2, \dots, N$ dengan N adalah panjang sinyal. Maka sinyal hasil normalisasi untuk menghilangkan komponen DC seperti Persamaan (1).

$$y(i) = x(i) - \frac{1}{N} \sum_{i=1}^N x(i) \quad (1)$$

Pada proses selanjutnya dilakukan proses seperti pada Persamaan (2).

$$y(i) = \frac{x(i)}{\max|x|} \quad (2)$$

Sinyal $y(i)$ adalah sinyal dengan amplitude -1 sampai +1 dengan demikian perbedaan rentang amplitude sinyal akibat perbedaan rekaman sinyal dapat dieliminasi.

2.2 Renyi Entropy

Renyi entropy (REN) adalah bentuk umum dari perhitungan entropy. Secara matematis REN dinyatakan dengan Persamaan (3)

$$S_q = \frac{1}{1-q} \log_2 \left(\sum_{i=1}^N p_i^q \right), \quad q \neq 1 \quad (3)$$

Dengan p_i adalah probabilitas kemunculan nilai sampel dalam sinyal $x(i)$. Biasanya untuk memudahkan perhitungan digunakan histogram sinyal. Secara praktis biasanya yang dinamakan REN hanya untuk $q = 2$.

Apabila digunakan order $q \neq 2$, maka Renyi entropy akan menghasilkan entropy yang berbeda. Jika $q = 0$, maka dinamakan dengan Hartley entropy, yang dirumuskan dengan Persamaan (4)

$$S_0 = \log_2 N \quad (4)$$

Dengan N adalah panjang sinyal. Sementara itu jika $q = 1$, maka akan didapat Shannon entropy seperti pada Persamaan (5).

$$S_1 = -\sum_{i=1}^N p_i \log_2 p_i \quad (5)$$

Pada penelitian ini digunakan Shannon entropy dan Renyi entropy dengan $q = 2-100$ sebagai ciri. Total ciri yang dihasilkan adalah 100 dan nantinya akan dikurangi tiap 10 ciri sampai mencapai 10 ciri dan kemudian dikurangi tiap satu ciri. Hal ini dilakukan untuk melihat efek jumlah ciri terhadap akurasi.

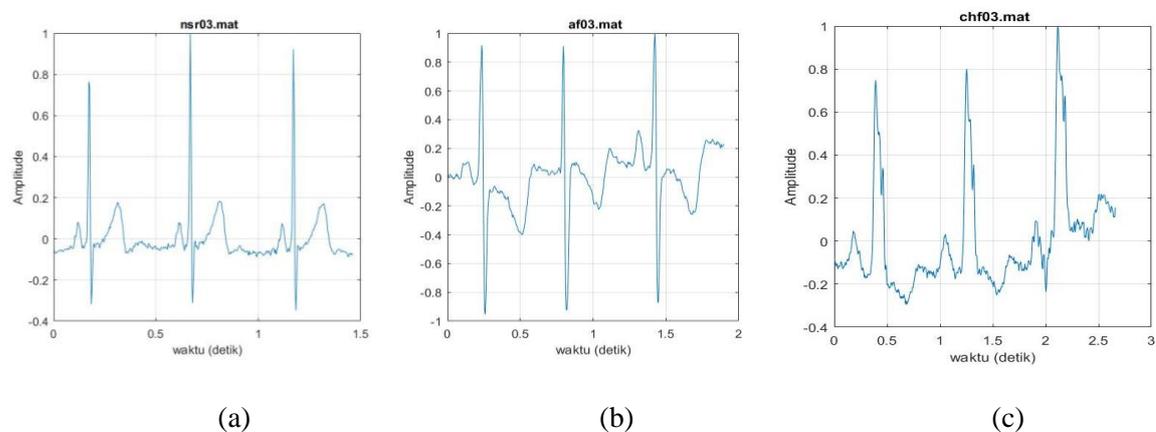
2.3 SVM dan N-fold Cross Validation

Support vector machine (SVM) adalah klassifier linear yang kemudian dikembangkan untuk menyelesaikan klasifikasi masalah non-linear. Konsep dasar dari SVM adalah mencari hyperplane terbaik untuk memisahkan dua kelas data [14]. Hyperplane adalah garis lurus atau bidang yang memisahkan data antar kelas. Hyperplane terbaik didapat dengan memaksimalkan margin antara dua objek dari kelas yang berbeda. Margin adalah jarak antara hyperplane dengan pola terdekat di tiap kelas data. Posisi terdekat antara pola tiap kelas dinamakan dengan support vector. Untuk memecahkan masalah non-linear digunakan metode yang dinamakan dengan kernel trick. Metode ini menggunakan bidang nonlinear untuk memisahkan data antar kelas yang mungkin tidak bisa dipisahkan oleh garis lurus. Pada penelitian ini digunakan Quadratic SVM dan Cubic SVM sebagai pembandingan linear SVM.

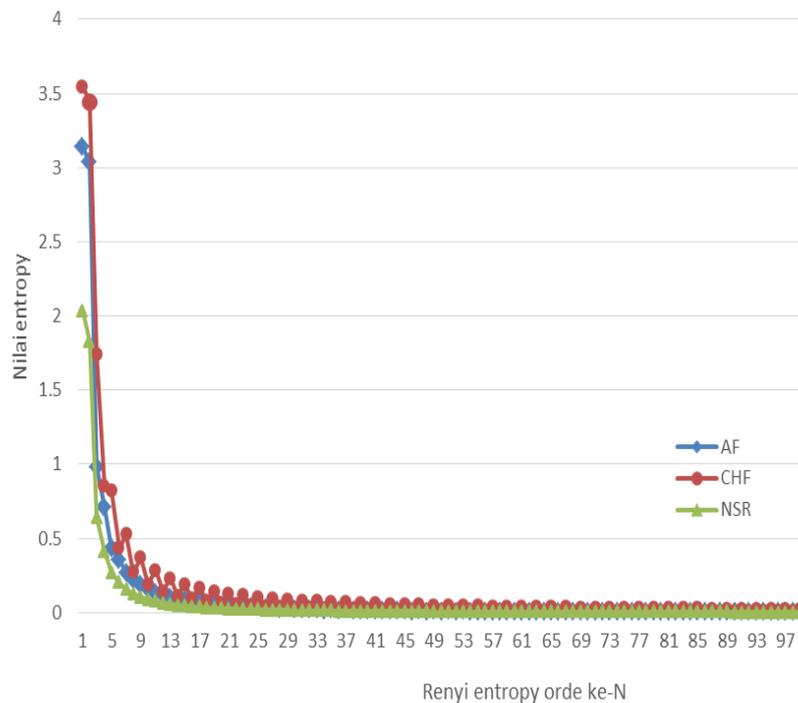
Karena SVM merupakan metode yang memerlukan pelatihan (supervised learning) maka pada penelitian ini digunakan N-fold cross validation (NFCV) untuk memisahkan data latih dan data uji. Pada NFCV, tiap kelas data dibagi menjadi N data set. N-1 data set digunakan sebagai data latih dan satu data set digunakan sebagai data uji. Proses tersebut diulang sampai N kali sehingga tiap data set pernah menjadi data uji [15]. Akurasi diambil dari rata-rata dari semua trial yang dilakukan [16]. Kelebihan metode ini dibandingkan dengan pembagian data latih dan data uji secara random adalah deviasi nilai akurasi yang lebih rendah. Parameter kinerja dari metode yang diusulkan adalah akurasi, yaitu jumlah data yang diklasifikasi dengan benar oleh system.

3. Hasil dan Diskusi

Contoh sinyal EKG untuk tiap kelas data seperti ditunjukkan pada Gambar 2. Sinyal EKG normal memiliki ritme yang teratur dengan sumbu mendatar yang relative rata, AF memiliki bentuk yang berbeda dengan bentuk sinyal EKG normal sedangkan CHF memiliki bentuk QRS yang khas karena menurunnya kemampuan jantung memompa darah [12]. Dengan melihat bentuk sinyal yang berbeda maka diharapkan Renyi entropy yang dihitung pada tiap sinyal mampu membedakan tiap kelas data.



Gambar 2. (a) Sinyal EKG normal (b) Sinyal EKG AF (c) Sinyal EKG CHF

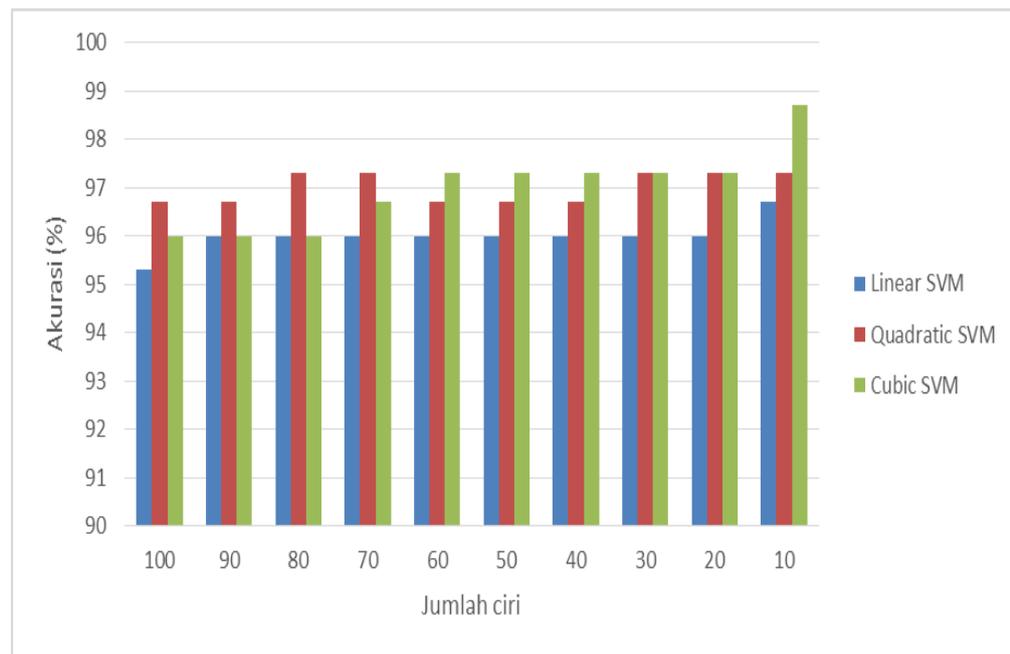


Gambar 4. Nilai rata-rata Renyi Entropy untuk $q = 1-100$ pada tiap kelas data

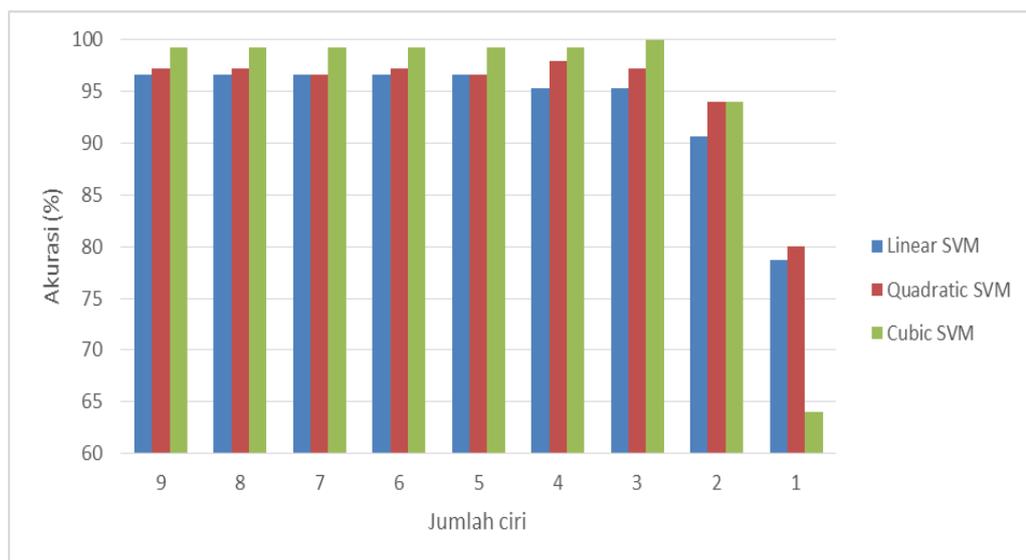
Gambar 4 menampilkan nilai REN rata-rata untuk tiap kelas data menggunakan $q = 1-100$. REN dengan $q = 1$ merupakan Shannon entropy. Terlihat bahwa nilai REN cukup besar sampai $q = 20$ -an dan selanjutnya cenderung mendatar karena nilai REN sangat rendah. Dari nilai REN yang cenderung menurun dapat diperkirakan bahwa tidak semua REN mempengaruhi akurasi klasifikasi. Untuk itu pada tahap selanjutnya dilakukan klasifikasi dan reduksi ciri untuk melihat akurasi yang dihasilkan.

Gambar 5 menampilkan akurasi klasifikasi sinyal EKG menggunakan tiga macam SVM, linear SVM, quadratic SVM dan Cubic SVM. Pengurangan jumlah ciri, dalam hal ini mengurangi jumlah orde q , menghasilkan akurasi yang relatif meningkat. Hal ini menunjukkan bahwa Renyi entropy dengan q tinggi menghasilkan nilai yang rendah dan cenderung sama untuk tiap kelas data seperti pada Gambar 4. Mengurangi ciri hingga mendekati 10 ciri meningkatkan akurasi utamanya menggunakan Cubic SVM sebagai classifier. Linear SVM menghasilkan akurasi 95.3% untuk penggunaan 100 ciri dan 96.7% untuk penggunaan 10 ciri. Quadratic SVM menghasilkan akurasi 96.7% untuk 100 ciri dan 97.3% untuk 10 ciri. Sementara itu Cubic SVM meningkatkan akurasinya dari 96% menggunakan 100 ciri menjadi 98.7%.

Pengurangan ciri lebih lanjut justru menurunkan akurasi seperti ditampilkan pada Gambar 6. Kecuali untuk Cubic SVM, akurasi tertinggi mencapai 100% untuk tiga ciri. Hal ini menunjukkan bahwa tiga nilai entropy pertama mampu membedakan tiga kelas data secara tepat. Penggunaan satu ciri saja dari ketiga ciri tersebut hanya menghasilkan akurasi tertinggi 86% untuk REN dengan $q = 3$, 87.3% untuk REN dengan $q = 2$ dan 80% untuk Shannon entropy. Hal ini membuktikan bahwa penggunaan tiga entropy sekaligus memberikan hasil yang lebih baik dibandingkan dengan penggunaan satu ciri saja.



Gambar 5 Akurasi untuk jumlah ciri 10 – 100



Gambar 6. Akurasi untuk jumlah ciri 1 – 9

Tabel 1 menampilkan perbandingan beberapa penelitian sebelumnya yang menggunakan data EKG yang sama. Penelitian pertama menggunakan dekomposisi paket wavelet level 5 sehingga menghasilkan 32 subband. Ciri yang digunakan adalah energy tiap subband dengan K-man clustering untuk melihat keterpisahan antar kelas data. Akurasi yang dihasilkan adalah 94.4% [17]. Pada penelitian berikutnya, dari 32 ciri yang dihasilkan dilakukan proses feature subset selection untuk memilih ciri yang memberikan ciri yang signifikan. Hasilnya didapat empat ciri terbaik dan akurasi meningkat menjadi 87.33% [18]. Dua penelitian berikutnya menggunakan Hjorth descriptor sebagai ciri, yaitu activity, mobility, dan complexity serta complexity pada orde yang lebih tinggi [19], [20]. Penggunaan tiga parameter Hjorth descriptor menghasilkan akurasi 100% sedangkan Complexity orde tinggi hanya menghasilkan akurasi 94%.

Tabel 1. Perbandingan dengan metode lain

Referensi	Metode	Jumlah ciri	Classifier	Akurasi
[17]	Dekomposisi paket wavelet	32	K-mean clustering	94.4%
[18]	Dekoposisi paket wavelet	4	K-NN	97.33%
[19]	Hjorth Descriptor	3	K-NN/MLP	100%
[20]	Complexity orde tinggi dari Hjorth Descriptor	5	K-NN/MLP	94%
Penelitian ini	Renyi entropy	3	Cubic SVM	100%

Metode Renyi entropy yang disulkan menghasilkan akurasi 100% menggunakan tiga ciri. Hal ini menunjukkan bahwa metode yang diusulkan mampu bersaing dengan metode yang lain dari sisi akurasi dan jumlah ciri yang digunakan. Kelebihan metode ini adalah hanya satu parameter yang digunakan tetapi dihitung pada orde q yang berbeda. Hal ini membuat kompleksitas komputasi menjadi sangat rendah. Sementara itu kelemahan dari metode ini adalah jumlah orde Renyi entropy yang digunakan sebagai ciri hanya bisa ditentukan lewat trial and error. Renyi entropy bisa dihitung pada orde negative seperti pada penelitian oleh Jelinek et al. yang menggunakan orde -5 sampai 5 [10]. Penelitian menggunakan Renyi entropy dengan rentang orde negative dan positif menjadi topic yang menarik penelitian selanjutnya.

4. Kesimpulan

Pada penelitian ini dipaparkan tentang penggunaan Renyi entropy pada beberapa orde untuk ekstraksi ciri sinyal EKG. Renyi entropy adalah salah satu matrik untuk melihat kompleksitas sinyal yang merupakan bentuk umum dari entropy lain seperti Shannon entropy. Penggunaan beberapa entropy sekaligus menghasilkan akurasi yang lebih tinggi dibandingkan dengan penggunaan satu nilai entropy saja. Dari pengujian dihasilkan akurasi 100% untuk tiga kelas data EKG menggunakan tiga ciri Renyi entropy. Metode yang diusulkan unggul dalam komputasi yang sederhana dan jumlah ciri yang lebih sedikit. Pengujian pada sinyal biomedis yang lain menarik untuk dilakukan untuk melihat kinerja Renyi entropy sebagai ciri.

Daftar Pustaka

- [1] W. J. Tompkins, "Electrocardiography," in *Biomedical Digital Signal Processing*, W. J. Tompkins, Ed. New Jersey: Prentice Hall, 2000, pp. 24–54.
- [2] T. A. M. Brosche, *The EKG Handbook*. Jones & Bartlett Publisher, 2010.
- [3] M. Costa, A. L. Goldberger, and C. K. Peng, "Multiscale entropy analysis of biological signals," *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, vol. 71, pp. 1–18, 2005.
- [4] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
- [5] A. Renyi, "On Measures of Entropy and Information," in *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability*, 1960, pp. 547–561.
- [6] C. Tsallis, "Possible generalization of Boltzman-Gibbs Statistics," *J. Stat. Phys.*, vol. 52,

no. 1/2, pp. 479–487, 1988.

- [7] O. A. Rosso et al., “Wavelet entropy : a new tool for analysis of short duration brain electrical signals,” *J. Neurosci. Methods*, vol. 105, pp. 65–75, 2001.
- [8] S. Z. Fan, J. R. Yeh, B. C. Chen, and J. S. Shieh, “Comparison of EEG Approximate Entropy and Complexity Measures of Depth of Anaesthesia During Inhalational General Anaesthesia,” *J. Med. Biol. Eng.*, vol. 31, no. 5, pp. 359–366, 2011.
- [9] A. Rizal, R. Hidayat, and H. A. Nugroho, “Entropy measurement as features extraction in automatic lung sound classification,” in *2017 International Conference on Control, Electronics, Renewable Energy and Communications (ICCREC)*, 2017, pp. 93–97.
- [10] H. F. Jelinek, D. J. Cornforth, M. P. Tarvainen, and N. T. Miloevic, “Multiscale renyi entropy and cardiac autonomic neuropathy,” *Proc. - 2015 20th Int. Conf. Control Syst. Comput. Sci. CSCS 2015*, pp. 545–547, 2015.
- [10a] J. Kurths, A. Voss, P. Saparin, A. Witt, H.J. Kleiner, N. Wessel, "Quantitative analysis of heart rate variability", *Chaos*, 5 (1995), pp. 88-94
- [11] Physionet.org, “ECG Database.” [Online]. Available: <http://physionet.org/physiobank/database/#ecg>.
- [12] J. L. Schuster, C. Spence, M. Jacobs, and A. Wilkonson, “Living with Advanced Congestive Heart Failure : A Guide for Family Caregivers,” 2002.
- [13] S. M. Al-khatib, N. A. Lapointe, and R. Chatterjee, “Treatment of Atrial,” *Comp. Eff. Rev.*, no. 119, 2013.
- [14] C. Cortes and V. Vapnik, “Support-vector networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.
- [15] 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.
- [16] R. Palaniappan, *Biological Signal Analysis*. Ventus Publishing ApS, 2010.
- [17] A. Rizal and V. Suryani, “Pengenalan Signal EKG Menggunakan Dekomposisi Paket Wavelet dan K-Means Clustering,” in *Proceeding Seminar Nasional Aplikasi Teknologi Inofrmasi 2008(SNATI 2008)*, 2008, vol. 2008, no. Snati, pp. 5–8.
- [18] A. Rizal, “Perbandingan Skema Dekomposisi Paket Wavelet untuk Pengenalan Sinyal EKG,” *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 4, no. 2, 2015.
- [19] A. Rizal and S. Hadiyoso, “ECG signal classification using Hjorth Descriptor,” in *2015 International Conference on Automation, Cognitive Science, Optics, Micro Electro-Mechanical System, and Information Technology (ICACOMIT)*, 2015, no. 2, pp. 87–90.
- [20] S. Hadiyoso and A. Rizal, “Electrocardiogram signal classification using higher-order complexity of hjorth descriptor,” *Adv. Sci. Lett.*, vol. 23, no. 5, pp. 3972–3974, 2017.