



PROSIDING

Seminar Nasional MIPA 2016

Naskah diseminarkan pada 5 November 2016 dan dipublikasikan pada
<http://conf.unnes.ac.id/index.php/mipa/mipa2016/schedConf/presentations>



Identifikasi Atrial Fibrillation pada Isyarat ECG Menggunakan Metode K-Nearest Neighbor (KNN)

Mohammad Rofi'i¹, Indah Soesanti², dan Hanung Adi Nugroho³

Departemen Teknik Elektro dan Teknologi Informasi

Fakultas Teknik, Universitas Gadjah Mada

Yogyakarta, Indonesia

email: rofii_s2te_12@mail.ugm.ac.id¹

Abstrak

Tujuan dari penelitian ini adalah untuk mengidentifikasi kelainan jantung atau aritmia berupa *atrial fibrillation* pada isyarat *electrocardiogram* (ECG). *Electrocardiogram* merupakan salah satu alat bantu yang paling banyak digunakan untuk melakukan diagnosis kelainan jantung. Data penelitian yang digunakan berasal dari Rumah Sakit Umum Daerah Tugurejo Semarang yang terdiri dari data pasien dengan kasus *atrial fibrillation* (AF) dan data ECG normal atau *normal sinus rhythm* (NSR). Data yang diambil dalam bentuk data cetak yang selanjutnya di lakukan scanning untuk mendapatkan data citra digital agar bisa diproses dengan komputer. Pada penelitian ini ada beberapa tahapan, diantaranya adalah pra-pengolahan, ekstraksi ciri, dan klasifikasi. Proses ekstraksi ciri berdasarkan ciri statistik (*mean*, *standar deviasi*, *kurtosis*, *varians*, *skewness*) yang selanjutnya diklasifikasi menggunakan metode KNN. Hasil yang didapatkan adalah akurasi sebesar 86%, sensitivitas 84,85%, dan spesifisitas 88,23%.

Abstract

The purpose of this research was to identify heart abnormalities or arrhythmias such as atrial fibrillation on the electrocardiogram signal (ECG). An electrocardiogram is one of the tools most widely used to diagnose heart abnormalities. The research data were collected from the General Hospital of Semarang Tugurejo that consists of data of patients with atrial fibrillation (AF) and the data is normal or normal sinus rhythm (NSR). The data are taken in the form of print data that is subsequently done scanning to obtain digital image data to be processed by computers. In this research, there are several stages, including pre-processing, feature extraction and classification. The process of feature extraction based on statistical value (*mean*, *standard deviation*, *kurtosis*, *variance*, *skewness*) were further classified using KNN. The results obtained are an accuracy of 86%, sensitivity of 84.85% and specificity of 88.23%..

Keywords: Arrhythmia; Atrial Fibrillation; normal synus rhythm; electrocardiogram.

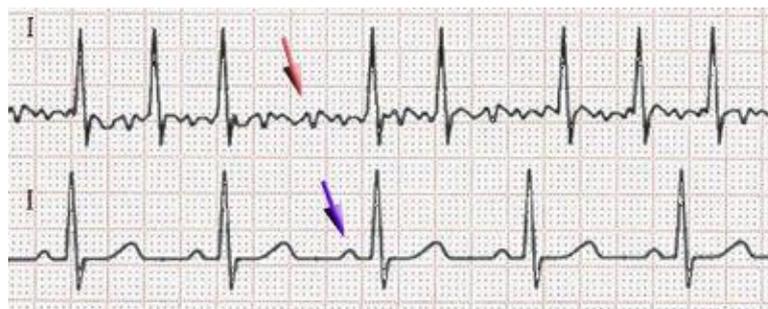
PENDAHULUAN

Jantung merupakan suatu organ otot berongga yang terletak di pusat dada. Fungsi utama jantung adalah menyediakan oksigen ke seluruh tubuh dan membersihkan tubuh dari hasil *metabolism* (karbon dioksida). Jantung melaksanakan fungsi tersebut dengan mengumpulkan darah yang kekurangan oksigen dari seluruh tubuh dan memompanya kedalam paru-paru, dimana darah akan mengambil oksigen dan membuang karbon dioksida, jantung kemudian mengumpulkan darah yang kaya oksigen dari paru-paru dan memompanya ke jaringan di seluruh tubuh (Sherwood, 2007). Fungsi vital yang diperankan oleh organ yang satu ini berpengaruh besar terhadap kondisi seseorang yang dapat dilihat dari isyarat yang dihasilkan oleh aktifitas kelistrikan jantung. *Electrocardiogram* merupakan salah satu rekaman isyarat fisiologi utama yang dihasilkan dari aktifitas kelistrikan jantung baik ketika fase *repolarisation* maupun *depolarisation*. ECG pertama kali diperkenalkan oleh Augustus

Desire Waller pada tahun 1887 (Waller, 1887). Sejak saat itu ECG menjadi salah satu isyarat yang penting dan paling sering digunakan untuk melakukan analisis terhadap aktifitas jantung (Rangayyan, 2002) dan telah menjadi *golden standard* untuk diagnosis aritmia jantung (Braunwald, 1997). Penelitian terhadap isyarat yang dihasilkan oleh jantung dengan menggunakan bantuan teknologi komputer sudah banyak dilakukan (Velic et al, 2013). Sehingga hal tersebut dapat dimanfaatkan oleh dokter ataupun tenaga medis dalam membaca dan menganalisis isyarat jantung yang dapat menampilkan abnormalitas aktifitas jantung. Hal itu merupakan salah satu media bantu dibidang kedokteran yang diharapkan oleh semua orang dalam rangka meningkatkan pelayanan kesehatan diseluruh dunia.

Kelainan bentuk isyarat ECG menunjukkan abnormalitas jantung yang biasanya disebut aritmia. Berdasarkan data dari badan kesehatan dunia (WHO) menempatkan abnormalitas jantung sebagai urutan teratas daftar penyebab utama kematian di seluruh dunia (Who, 2013). Aritmia merupakan istilah umum untuk setiap irama jantung yang berbeda dari irama sinus normal jantung atau *normal sinus rhythm* (NSR) (Sandoe, 1991). Aritmia yang paling umum dan yang sering dijumpai adalah *atrial fibrillation* (AF) (Chinthia et al, 2009). *Atrial fibrillation* memang telah menjadi jenis aritmia yang paling umum dan paling banyak terjadi dengan prevalensi meningkat setiap tahunnya. AF diperkirakan berpengaruh terhadap 0,5% sampai 1% dari penduduk Amerika Serikat, atau setara dengan sekitar 2,3 juta orang (Lainscak, 2008). AF juga merupakan kelainan jantung yang paling banyak pada kalangan orang tua, dengan usia rata-rata dari mereka yang terkena dampak ini berada pada sekitar 75 tahun. Prevalensi AF meningkat mencapai 8% pada orang Amerika Serikat yang berusia diatas 80 tahun. Faktor resiko yang berkontribusi terhadap fibrilasi atrium termasuk obesitas, penyakit arteri koroner, diabetes, gagal jantung, hipertensi, *hyperthyroidism*, *myocardial infarction*, dan *alcoholism* (Sunil et al, 2009).

Identifikasi dan analisis terhadap isyarat ECG telah menarik banyak perhatian untuk melakukan penelitian terkait abnormalitas jantung. Dari berbagai jenis aritmia jantung yang paling umum adalah fibrilasi atrium. Penerapan beberapa metode analisis terhadap isyarat ECG umumnya dilakukan pada domain waktu (*time domain*), kawasan frekuensi (*frequency domain*), kawasan spasial (*spatial domain*), dan *non linear statistic*.



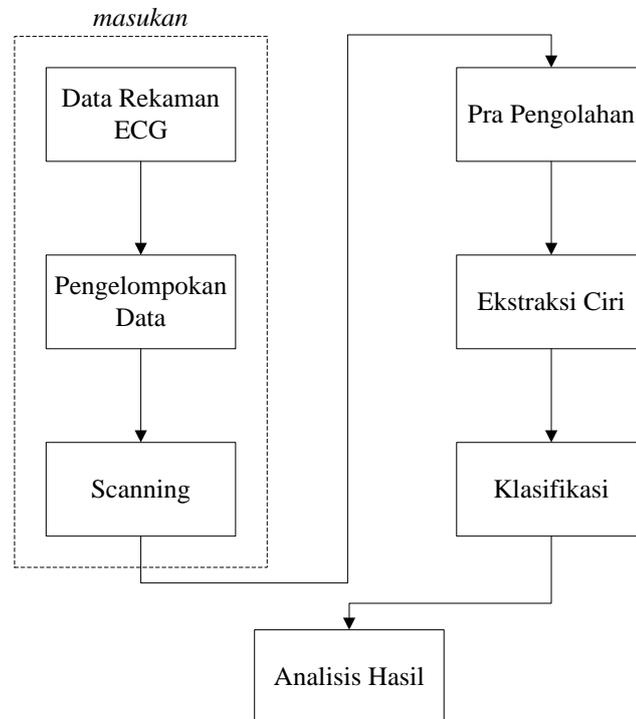
Gambar 1. Gelombang P yang diganti dengan Gelombang fibrillatory

Analisis terhadap interval R-R merupakan salah satu cara yang sering digunakan untuk mendeteksi abnormalitas jantung (Mark et al., 1991), karena AF digambarkan oleh ketidakteraturan interval R-R dan tidak adanya gelombang P normal yang telah diganti dengan gelombang fibrillatory (Gambar 1) yang bervariasi dalam ukuran, bentuk dan timing (Fuster et al., 2006).

METODE

Bahan yang digunakan dalam penelitian ini adalah data fisik electrocardiogram (masih berbentuk data cetak di atas kertas grafik) untuk dua jenis kondisi jantung yang diambil dari

Rumah Sakit Umum Daerah Tugurejo Semarang yang selanjutnya dilakukan scanning untuk mendapatkan data digital berupa citra dua dimensi. Jenis data yang diambil adalah data ECG normal berjumlah 20 data dan *Atrial Fibrillation* (AF) berjumlah 30 data, sehingga jumlah keseluruhan data ada 50 data. Pengolahan isyarat ECG dalam penelitian ini dilakukan melalui beberapa tahapan, diantaranya adalah pra pengolahan, ekstraksi ciri, dan klasifikasi. Tahapan tersebut dapat dilihat pada Gambar 2.



Gambar 2. Tahapan penelitian isyarat EKG

Tahap pra-pengolahan dilakukan untuk mendapatkan data yang lebih baik dari data sebelumnya. Pada tahap ini selain dilakukan segmentasi terhadap citra isyarat ECG juga dilakukan transformasi dari kawasan spasial (*spacial domain*) ke kawasan waktu (*time domain*), sehingga didapatkan isyarat ECG dalam domain waktu. Selanjutnya dilakukan proses ekstraksi ciri dari isyarat tersebut. Pada tahapan ekstraksi ciri terlebih dahulu dicari isyarat PSD dengan metode welch. Ekstraksi ciri statistik diambil dari isyarat PSD ternormalisasi. Ciri statistik yang digunakan adalah :

- *Mean* (μ)
Mean merupakan nilai rata-rata dari beberapa buah data

$$\mu = \frac{1}{N} \sum_{i=0}^{N-1} X_i$$

- Standar deviasi (σ)
Standar deviasi atau simpangan baku merupakan akar kuadrat varian

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (x_i - \mu)^2}$$

- Kurtosis (α^4)

Menggambarkan keruncingan (*peakedness*) atau kerataan (*flatness*) suatu distribusi data dibandingkan dengan distribusi normal

$$\alpha^4 = E\{y^4\} - 3(E\{y^2\})^2$$

- Varians (σ^2)
Varians atau ragam suatu peubah acak (distribusi probabilitas) adalah ukuran bagi persebaran atau dispersi data. Yang diukur adalah seberapa jauh data tersebar disekitar rerata

$$\sigma^2 = var(x) = \frac{1}{N-1} \sum_{i=0}^{N-1} (x_i - \mu)^2$$

- *Skewness* (α^3)
Skewnees merupakan pengukuran tingkat ketidaksimetrisan (kecondongan) sebaran data disekitar rata-ratanya

$$\alpha^3 = \frac{E(x - \mu)^3}{\sigma^3}$$

Setelah didapatkan ciri statistik tersebut selanjutnya diklasifikasi dengan metode *k-nearest neighbor* (KNN). Secara sederhana algoritma ini bekerja dengan membandingkan jarak data masukan dengan sejumlah k data pelatihan yang paling dekat (Pang, 2006). Pengujian terhadap metode klasifikasi yang diterapkan dalam penelitian ini dengan cara menghitung nilai akurasi, sensitivitas, dan spesifisitas.

$$Akurasi = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Sensitivitas = \frac{TP}{TP + FN}$$

$$Spesifisitas = \frac{TN}{FP + TN}$$

Keterangan :

TP = True Positive

FP = False Positive

TN = True Negative

FN = False Negative

HASIL DAN PEMBAHASAN

Pra-pengolahan

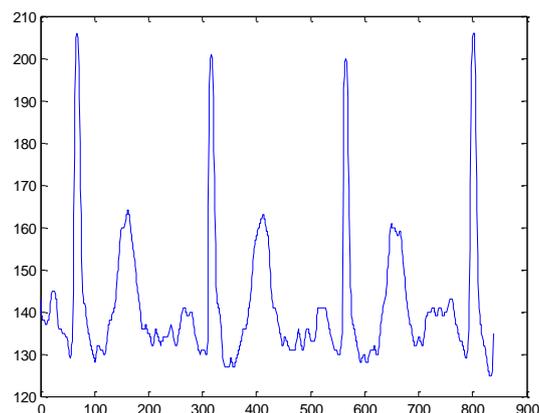
Pra-pengolahan dilakukan untuk mendapatkan data yang lebih baik dari data sebelumnya (data asli). Dengan cara memanipulasi parameter-parameter yang terdapat pada data tersebut. Sehingga didapatkan bentuk yang lebih cocok terhadap nilai-nilai yang diinginkan. Operasi pra pengolahan yang dilakukan dimulai dengan pengolahan citra rekaman ECG hasil scanning sebelumnya, sehingga diperoleh citra yang memiliki ukuran dan dimensi serta sebaran warna (nilai piksel) sesuai dengan yang diinginkan sehingga dapat digunakan untuk proses selanjutnya. Tahap pertama yang dilakukan adalah *cropping* citra

atau memotong suatu bagian dari citra sehingga diperoleh citra berukuran lebih kecil. Langkah selanjutnya adalah segmentasi.



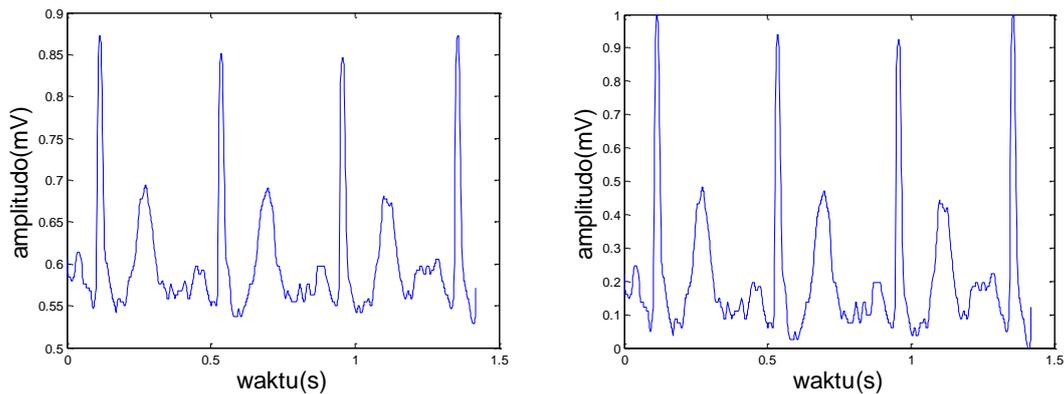
Gambar 3. (a) Citra grayscale ECG (b) Hasil segmentasi citra ECG

Segmentasi dilakukan untuk mendapatkan citra yang hanya terdiri dari objek yang diinginkan dan membuang objek lain yang dapat mengganggu proses selanjutnya seperti pada Gambar 3. pada tahap ini sebelumnya citra diubah kedalam bentuk citra keabuan atau *grayscale*, kemudian dilakukan pemfilteran dengan menggunakan *median filter*. Citra hasil filter diubah ke dalam citra biner yang hanya terdiri dari warna hitam dan putih saja. Setelah didapatkan citra biner kemudian dilakukan operasi morfologi *closing*. Operasi morfologi ini berfungsi untuk menghilangkan bagian detail yang terlihat gelap dan menyisakan bagian terang yang tidak mengganggu. Langkah selanjutnya adalah melakukan transformasi kawasan spasial ke kawasan waktu.



Gambar 4. Hasil konversi ke satu dimensi

Proses ini mengubah citra dua dimensi menjadi satu dimensi (Gambar 4). Hasil yang didapat tinggi citra dari isyarat menjadi penggambaran amplitudo isyarat dan pergeseran pembacaan menjadi penunjuk waktu, jika citra mempunyai resolusi 600 dpi maka setiap 25,4 mm (1 *inch*) terdapat 600 piksel. Dengan demikian lebar citra (x) dalam ukuran mm adalah $\frac{x}{600} \times 25,4 \text{ mm}$ dan untuk tinggi isyarat adalah $\frac{y}{600} \times 25,4 \text{ mm}$. Karena besaran yang ada pada isyarat adalah amplitudo dan waktu maka sumbu y (tinggi isyarat) diubah kedalam satuan mV dimana dalam ECG 1 mm menunjukkan 0,1 mV sehingga amplitudo isyarat adalah $y \times 0,1 \text{ mV}$. Sedangkan pada pengukuran sumbu x (lebar isyarat) diubah ke dalam satuan waktu (detik). Laju perekaman yang dipakai adalah 25 mm/detik, maka 1 mm setara dengan 0,04 detik, sehingga waktu isyarat (t) diperoleh dengan mengalikan $x \times 0,04 \text{ detik}$.



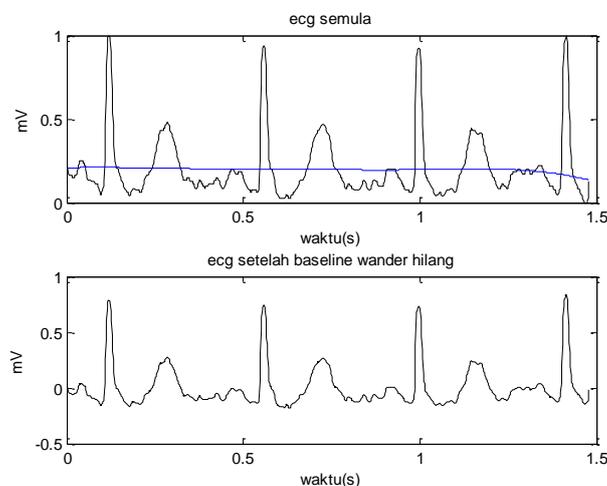
Gambar 5. Isyarat dalam kawasan waktu

Setelah berubah dalam kawasan waktu, maka kemudian dilakukan normalisasi (Gambar 5). Isyarat hasil transformasi dinormalisasi dengan tujuan menyederhanakan bentuk isyarat agar pengolahan lebih efektif, seperti pada persamaan di bawah ini :

$$x'_n = x_n - \frac{x_{min} + x_{max}}{2}$$

$$y_n = \frac{x'_n}{x'_{max}}$$

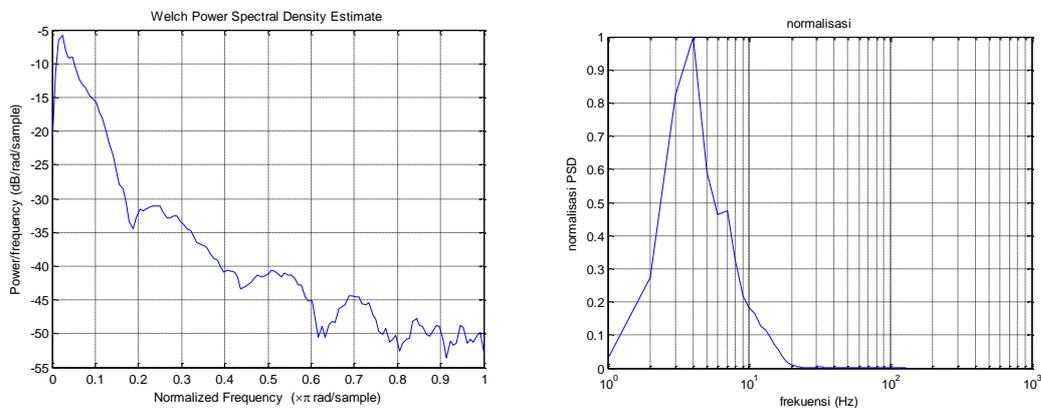
Langkah terakhir pada tahap proses pra pengolahan adalah *baseline wander removal*. Penghapusan derau *baseline wander* dapat dilakukan dengan menggunakan transformasi wavelet diskrit. Prinsip dasar dari transformasi wavelet diskrit adalah bagaimana cara mendapatkan representasi waktu dan skala dari sebuah isyarat menggunakan teknik pemfilteran digital dan operasi *sub-sampling*. Isyarat ECG pertama-tama dilewatkan pada rangkaian filter *high-pass* dan *low-pass*, kemudian setengah dari masing-masing keluaran diambil dengan sampel melalui operasi *sub-sampling*. Proses ini disebut sebagai proses dekomposisi satu tingkat. Keluaran filter *low-pass* digunakan sebagai masukan proses dekomposisi tingkat berikutnya. Proses ini diulang sampai tingkat proses dekomposisi yang diinginkan. Pada penelitian ini isyarat ECG didekomposisi hingga tingkat 11, sehingga diperoleh komponen frekuensi rendah yang menjadi penyebab derau *baseline wander*. Hal ini dapat dilihat pada Gambar 6 dibawah ini.



Gambar 6. Isyarat tanpa derau baseline wander

Ekstraksi ciri

Proses ekstraksi ciri dilakukan bertujuan untuk mendapatkan ciri dari masing-masing jenis isyarat ECG, yaitu isyarat ECG normal atau *normal sinus rhythm* (NSR) dan *atrial fibrillation* (AF). Ciri yang diambil merupakan nilai statistik isyarat PSD (power spectral density) ternormalisasi dari isyarat ECG. Nilai statistik yang digunakan adalah *mean*, *standard deviation*, *kurtosis*, *variance*, dan *skewness*.



Gambar 7. Isyarat power spectral density (PSD)

Setelah didapatkan ciri statistiknya maka langkah selanjutnya adalah proses klasifikasi dengan metode KNN.

Dari grafik isyarat diatas dapat dicari ciri statistiknya, sehingga diperoleh

Mean	= 0,0402
Std deviasi	= 0,1427
Varian	= 0,0204
Kurtosis	= 26,6617
Skewnes	= 4,6916

Klasifikasi

Klasifikasi merupakan proses penemuan model atau fungsi yang menggambarkan dan membedakan kelas data atau konsep yang bertujuan agar bisa digunakan untuk memprediksi kelas dari objek yang label kelasnya tidak diketahui (Han J., et al, 2006). Metode klasifikasi pada penelitian ini adalah k-nearest neighbor (KNN). Prinsip kerja KNN adalah mencari jarak terdekat antara data yang akan dievaluasi (data *testing*) dengan dengan k tetangga (*neighbor*) terdekatnya dalam data pelatihan (data *training*). Untuk mengukur kedekatan antara data baru dan data lama atau data training menggunakan beberapa metode pengukuran jarak, diantaranya adalah *euclidean distance* dengan persamaan berikut :

$$d(x, y) = \sqrt{\sum_{i=1}^k (x_i - y_i)^2}$$

Dimana $x = x_1, x_2, \dots, x_k$ dan $y = y_1, y_2, \dots, y_k$ mewakili k nilai atribut dari data testing dan data training. Manhattan distance atau dikenal city block distance merepresentasikan jarak antara dua objek secara absolut, dengan persamaan sebagai berikut :

$$d(x, y) = \sum_{i=1}^k |x_i - y_i|$$

Chebyshev distance disebut juga maximum value distance yaitu menghitung max dari besaran absolut perbedaan antara koordinat dari sepasang objek. Chebyshev distance direpresentasikan dalam persamaan seperti berikut:

$$d(x, y) = \max_i |x_i - y_i|$$

Evaluasi terhadap metode klasifikasi untuk mendapatkan hasil estimasi akurasi pada penelitian ini menggunakan metode cross validation dengan fold=5. Sehingga jumlah dari keseluruhan data akan dibagi kedalam 5 bagian yang sama. Jika terdapat 50 instances (sampel) maka akan terdapat 5 blok dengan masing-masing blok terdiri dari 10 instance. Jika kelima blok tersebut adalah A,B,C,D,dan E, maka pengujian dilakukan dengan cara sebagai berikut :

Step 1 : training memakai A,B,C,D dan testing memakai E → akurasi *a*

Step 2 : training memakai A,B,C,E dan testing memakai D → akurasi *b*

Step 3 : training memakai A,B,D,E dan testing memakai C → akurasi *c*

Step 4 : training memakai A,C,D,E dan testing memakai B → akurasi *d*

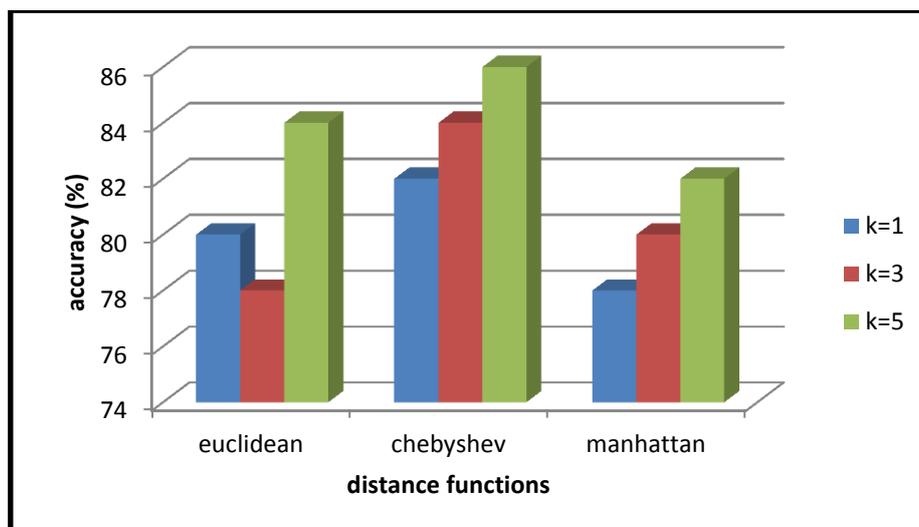
Step 5 : training memakai B,C,D,E dan testing memakai A → akurasi *e*

Sehingga rata-rata akurasi didapatkan $(a+b+c+d+e)/5$.

Tabel 1. Hasil perhitungan akurasi

Accuracy			
	K=1	K=3	K=5
Euclidean	80 %	78 %	84 %
Chebyshev	82 %	84 %	86 %
Manhattan	78 %	80 %	82 %

Berdasarkan tabel diatas tampak bahwa akurasi tertinggi pada saat k= 5 dan perhitungan jarak dengan menggunakan metode *chebyshev distance* yaitu sebesar 86%.



Gambar 8. Pengaruh nilai k dan jenis *distance function* terhadap akurasi

KESIMPULAN

Berdasarkan pengujian yang telah dilakukan terkait identifikasi atrial fibrillation (AF) dapat dilakukan dengan beberapa cara diantaranya adalah dengan metode pengolahan isyarat dengan bantuan komputer. Dari hasil diatas dapat disimpulkan bahwa nilai akurasi tertinggi klasifikasi dengan metode KNN didapatkan pada saat $k=5$ dan dengan menggunakan metode pengukuran jarak chebyshev distance dengan akurasi sebesar 86%, Sensitivitas 84,85%, dan spesifisitas 88,23%.

DAFTAR PUSTAKA

- Who (2013). "The 10 leading causes of death in the world". <http://www.who.int/mediacentre/factsheets/fs310/en/> diakses 13 oktober 2016
- Sherwood. (2007) . "Human Physiology: From Cells to Systems". 6th Edition Thomas Brooks/Cole, 2007. Belmont, CA
- Waller, A. (1887). "A demonstration on man of electromotive changes accompanying the heart's beat". *J Physiol (Lond)* 1887;8:229-34
- Rangayyan, R. M. (2002). "Biomedical Signal Analysis", Wiley-Interscience, John Wiley & SONS, INC., 2002
- Braunwald, E. (1997). "Heart Disease: A Textbook of Cardiovascular Medicine", Fifth Edition, p. 108, Philadelphia, W.B. Saunders Co., 1997. ISBN 0-7216-5666-8
- Velic, M., et al. (2013). "Computer aided ECG analysis — State of the art and upcoming challenges", *EUROCON, 2013 IEEE*, Pages: 1778 - 1784, DOI: 10.1109/EUROCON.2013.6625218, 2013
- Sandoe, E., Et al. (1991). "Arrhythmia : a Guide to Clinical Electrocardiology", Verlags GmbH, Bingen, 1991 (Chapter 3).
- Cynthia, A., et al. (2009). "Clinical, Economic, and Quality of Life Impact of Atrial Fibrillation FCCP, BCPS." *Manag Care Pharm.* 2009;15(6-b)(Suppl):S4-S9
- Lainscak, M., et al. (2008). "Atrial fibrillation in chronic non-cardiac disease: where do we stand?" *Int J Cardiol* 2008;128(3):311-315
- Sunil, T., et al. (2009, november). "Atrial fibrillation: Mechanistic insights and treatment options" *European Journal of Internal Medicine*, Volume 20, Issue 7, November 2009, Pages 672-681
- Artis. S., et al. (1991). "Detection of atrial fibrillation using artificial neural networks", *Computers in Cardiology*, 1991, pp 173-176
- Fuster. V., et al. (2006). "Guidelines for the Management of Patients With Atrial Fibrillation", *Circulation* vol 114, 2006, pp 700-752. ACC/AHA/ESC 2006
- Pang-N., T., et al. (2006). "Introduction to Data Mining", 1st penyunt., Boston: Pearson Addison Wesley
- Han, J., et al. (2006). "Data Mining Concept and Tehniques". San Fransisco : Morgan Kauffman.