

## ELECTROCARDIOGRAM ARRHYTHMIA CLASSIFICATION SYSTEM USING SUPPORT VECTOR MACHINE BASED FUZZY LOGIC

Sugiyanto, Tutuk Indriyani, and Muhammad Heru Firmansyah

Department of Informatics Engineering, Faculty of Information Technology,  
Adhi Tama Institute of Technology Surabaya  
Arief Rachman Hakim 100, Surabaya, 60117, Indonesia

E-mail: [sugianto@itats.ac.id](mailto:sugianto@itats.ac.id), [tutuk@itats.ac.id](mailto:tutuk@itats.ac.id), [herufirmansyah93@gmail.com](mailto:herufirmansyah93@gmail.com)

### Abstract

Arrhythmia is a cardiovascular disease that can be diagnosed by doctors using an electrocardiogram (ECG). The information contained on the ECG is used by doctors to analyze the electrical activity of the heart and determine the type of arrhythmia suffered by the patient. In this study, ECG arrhythmia classification process was performed using Support Vector Machine based fuzzy logic. In the proposed method, fuzzy membership functions are used to cope with data that are not classifiable in the method of Support Vector Machine (SVM) one-against-one. An early stage of the data processing is the baseline wander removal process on the original ECG signal using Transformation Wavelet Discrete (TWD). Afterwards then the ECG signal is cleaned from the baseline wander segmented into units beat. The next stage is to look for six features of the beat. Every single beat is classified using SVM method based fuzzy logic. Results from this study show that ECG arrhythmia classification using proposed method (SVM based fuzzy logic) gives better results than original SVM method. ECG arrhythmia classification using SVM method based fuzzy logic forms an average value of accuracy level, sensitivity level, and specificity level of 93.5%, 93.5%, and 98.7% respectively. ECG arrhythmia classification using only SVM method forms an average value accuracy level, sensitivity level, and specificity level of 91.83%, 91.83%, and 98.36% respectively.

**Keywords:** *arrhythmia classification, ECG, fuzzy logic, heart rate, Support Vector Machine*

### Abstrak

Aritmia adalah penyakit kardiovaskular yang dapat didiagnosis dokter menggunakan elektrokardiogram (EKG). Informasi yang terdapat di EKG digunakan oleh dokter untuk menganalisis aktivitas elektrik jantung dan menentukan jenis aritmia yang diderita oleh pasien. Dalam penelitian ini, proses klasifikasi aritmia EKG dilakukan dengan menggunakan Support Vector Machine berbasis fuzzy logic. Pada metode yang diusulkan, fungsi keanggotaan fuzzy digunakan untuk mengatasi dengan data yang tidak dapat diklasifikasikan dalam metode Support Vector Machine (SVM) satu-terhadap-satu. Tahap awal pengolahan data adalah proses *baseline wander removal* pada sinyal EKG asli menggunakan Transformasi Wavelet Diskrit (TWD), dan kemudian sinyal EKG bersih dari *baseline wander* tersegmentasi ke unit denyut. Tahap berikutnya adalah untuk mencari enam fitur dari denyut, dan setiap denyut tunggal diklasifikasikan menggunakan metode SVM berbasis fuzzy logic. Hasil dari penelitian menunjukkan bahwa klasifikasi aritmia EKG menggunakan metode yang diusulkan (SVM berdasarkan logika fuzzy) memberikan hasil yang lebih baik daripada metode SVM asli. Klasifikasi aritmia EKG menggunakan metode SVM berbasis logika fuzzy membentuk nilai rata-rata tingkat akurasi, tingkat sensitivitas, dan tingkat spesifisitas 93,5%, 93,5%, dan 98,7%. Klasifikasi aritmia EKG menggunakan metode SVM asli hanya membentuk tingkat rata-rata nilai akurasi, tingkat sensitivitas, dan tingkat spesifisitas 91,83%, 91,83%, dan 98,36%.

**Kata Kunci:** *klasifikasi aritmia, ECG, logika fuzzy, denyut jantung, Support Vector Machine*

### 1. Introduction

Arrhythmias are disorders of the heart in the form of interference on the frequency, irregularity, place of origin pulse or conduction of electrical impulses in the heart [1]. Arrhythmia is a dangerous disease, so patients need immediate treatment and

therapy regularly to prevent a worsening condition [2]. In general, the diagnosis of arrhythmias can only be done by a cardiologist. Along with the development of science and technology, many researchers doing research on the diagnosis of arrhythmias to find a system that can classify arrhythmias more accurately.

Many researchers have previously raised the topic of arrhythmia classification. Srivastava et al. [3] create an arrhythmia classification system using Fuzzy Sugeno method. The proposed system is able to categorize an ECG wave to one of thirteen types of arrhythmia. Results from the classification system are consistent with the results of testing by the cardiologist as many as 91 of the 105 data.

Moavenian and Khorrami [4] conducted a comparison using the method of Support Vector Machine (SVM) and Artificial Neural Network (ANN) to classify the six types of arrhythmia among others Right Bundle Branch Block (RBBB), Left Bundle Branch Block (LBBB), Premature Ventricular Beat (PVB), Paced Beat (PB) Premature Atrial Beat (PAB), and Fusion of Paced and normal Beat (FB). The testing process using three assessment criteria includes performance training, performance testing and training time. The results show that SVM is superior in performance training and training time while ANN is superior in performance testing.

Castillo et al. [5] did a comparison between three methods of classification of arrhythmias include Fuzzy K-Nearest Neighbours, Multi-Layer Perceptron with Back Propagation Gradient Descent with momentum, and Multi-Layer Perceptron with Back Propagation Gradient Scale Conjugated. Each method resulted considerable accuracy. After the arrhythmia classification system combined with Fuzzy Mamdani method, it rise the accuracy up to 98%.

From several previous studies, the authors found a solution to make the arrhythmia classification system using SVM method that has proven better for classifying arrhythmias, and combined with fuzzy methods to address data that is not classified. One of the electrocardiogram signal noises that often arise is the Baseline Wander, a condition in which an electrocardiogram signal have movement at low frequencies irregular [6]. So that the system can perform accurate classification electrocardiogram signal, the signal must be free from noise. The solution resulted in higher levels of accuracy.

## 2. Methods

### Electrocardiogram

Electrocardiogram (ECG) is a representation of a signal generated by the electrical activity of the heart muscle [7]. ECG signal is recorded using an electrocardiograph device. Electrocardiograph devices are medical devices used by patients to measure the electrical activity of the heart by measuring the difference biopotential from the outside

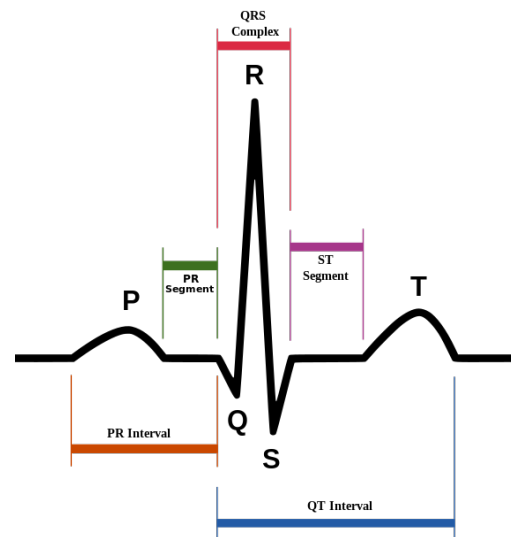


Figure 1. Schematic representation of normal ECG

into the body. In the medical field, electrocardiograph is used to diagnose some types of diseases related to the heart of which is a heart attack, disease/heart conditions, ischemia, hypertension (high blood pressure). One-piece ECG describes the condition of normal heartbeat consists of one P wave, one QRS complex, and one T wave. Figure 1 shows a schematic representation of normal ECG.

### Arrhythmia

Arrhythmia is a change in the frequency of heart rhythm caused by abnormal electrolyte conduction or automatically [1]. Heart rhythm disorder is not confined to the irregularity of the heart rate but also includes the rate and conduction disturbances. Heart rhythm disorder is caused by lack of oxygenated blood supply to the heart muscle, usually arises throbbing sensation that is too slow, too fast, or irregular throbbing pulsations [8]. Based on the type of beat, Association for the Advancement of Medical Instrumentation (AAMI) classifies the arrhythmia into 15 types of beat [9]. This study will classify the six types of disorder's arrhythmias selected from disorders that often appear in the dataset MIT-BIH [10] among other's Normal Beat (NB), Premature Ventricular Contraction (PVC), Paced Beat (PB), Left Bundle Branch Block Beat (LBBB), Right Bundle Branch Block Beat (RBBB), and Atrial Premature Beat (APB).

### Support Vector Machine

Support Vector Machine (SVM) is a linear classifier method with the predetermined feature set [11]. SVM delivers maximum results while using fewer data training and there is no overlap bet-

ween the existing classes [12]. SVM looking for a hyperplane with the largest margin is called Maximum Marginal Hyperplane (MMH) to separate the existing class.  $D(x, w, b)$  is a decision function to determine the MMH can be expressed by equation(1).

$$D(x, w, b) = W^T \cdot X + b = \sum_{i=1}^n W_i x_i + b \quad (1)$$

To optimize the equation(1) in order to obtain a maximum margin, quadratic programming optimization is performed using equation(2).

$$L_d(a) = -0.5a^2 H a + f^T a \quad (2)$$

Equation(2) has a constraint such as equation(3), and equation(4).

$$y^T a = 0 \quad (3)$$

$$a_i \geq 0, i = 1, \dots, l, \quad (4)$$

where  $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_l]^T$ , H is the Hessian matrix notation for  $H_{ij} = y_i y_j x_i^T x_j$ , f is a (l, 1) vector, and  $f = [1 \ 1 \ \dots \ 1]^T$ . With regard to equation (1), then we used equation(5) and (6) to obtain the value of w, and b are optimal.

$$W_o = \sum_{i=1}^l a_{oi} y_i x_i \quad (5)$$

$$b_0 = \frac{1}{N_{sv}} \left( \sum_{s=1}^{N_{sv}} \left( \frac{1}{y_s} - x_s^T w_o \right) \right) \quad (6)$$

where equation(6) is only used on data that has a support vector ( $\alpha > 0$ ).

### Support Vector Machine Based Fuzzy Logic

Support Vector Machine (SVM) based fuzzy logic is the development of Support Vector Machine methods to overcome the problems of multiclass. In each class, there will be defined a polyhedral pyramidal decision membership function using function obtained from SVM method for a class pair. SVM method will look for a hyperplane with the largest margin called Maximum Marginal Hyperplane (MMH). The distance between the hyperplane with a side of the margin is equal to the distance between the hyperplane with margin on the other side. If the classification is done in pairs, decision function for class i and j formulated in equation(7).

$$D_{ij}(x) = w_{ij}^T x + b_{ij} \quad (7)$$

SVM based fuzzy logic used membership functions to classify areas that cannot be classified by the decision function. Equation(8) shows the membership function  $m_{ij}$ .

$$m_{ij}(x) = \begin{cases} 1, & \text{for } D_{ij}(x) \geq 1, \\ D_{ij}(x) & \text{for the other} \end{cases} \quad (8)$$

By using  $m_{ij}(x)$ , the membership function x to the class i can be defined using the operator equation(9).

$$m_i(x) = \min_{j=1, \dots, n} m_{ij}(x) \quad (9)$$

Equation(8) and (9) are equivalent to equation (10).

$$m_i(x) = \min(1, \min_{j \neq i, j=1, \dots, n} D_{ij}(x)) \quad (10)$$

To simplify the calculations, equation(10) can be written into the equation(11).

$$m_i(x) = \min_{j \neq i, j=1, \dots, n} D_{ij}(x) \quad (11)$$

Furthermore, the data is classified based on the highest membership value according to the equation (12).

$$\text{Data class } x = \max_{i=1, \dots, n} m_i(x) \quad (12)$$

### ECG Dataset

The dataset used in this study is the MIT-BIH arrhythmia database published by PhysioNet [10]. The source of the ECG recording consists of 4000 Holter recordings originating from Beth Israel Hospital Arrhythmia Laboratory between 1975 and 1979. ECG randomly selected each from 48 the data recording ECG signals with a duration of 30 minutes. Examples of ECG signal representation can be seen in Figure 2.

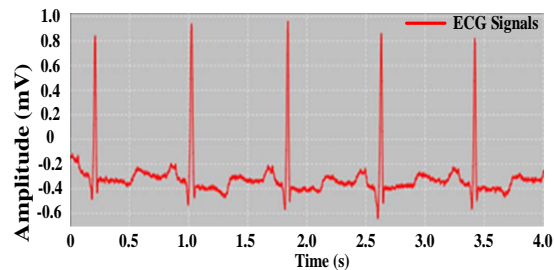


Figure 2. ECG signal representation

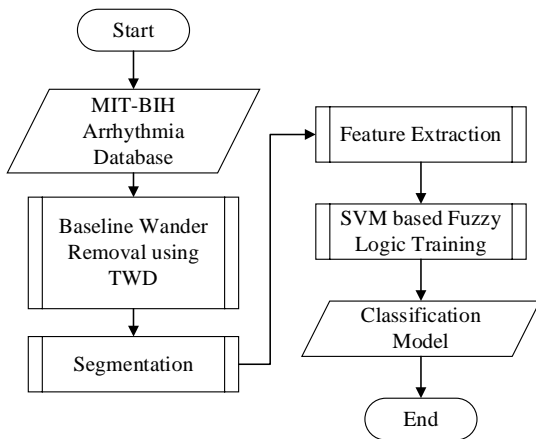


Figure 3. Training System Flowchart

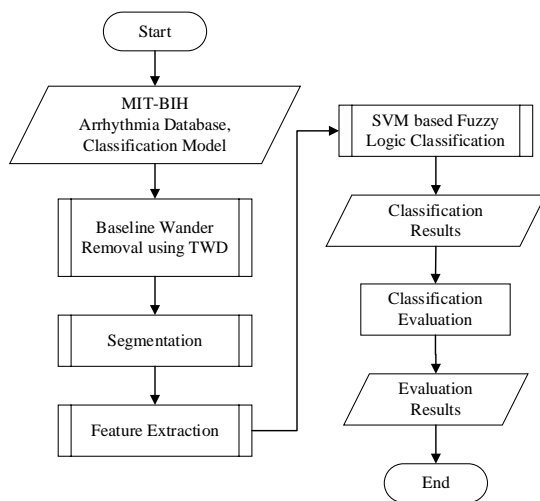


Figure 4. Testing System Flowchart

### System Flowchart

System flowchart is divided into two parts, namely the training system flowchart, and testing system flowchart. Training system doing the learning from the training data to find a classification model. The early stage of ECG data processing is the baseline wander removal process on the original ECG signal using Discrete Wavelet Transformation (TWD). The ECG signal is clean from the baseline wander will be segmented into units beat. The next stage is to find six features of the beat, and every single beat is classified using SVM method based fuzzy logic. Researchers added fuzzy membership function in SVM method to solve the problem of classified data when using Support Vector Machine (SVM) one-against-one. Figure 3 shows the training system flowchart.

The testing system created a prediction based on testing data processing to get the classification results, and continued by analyzing the cla-

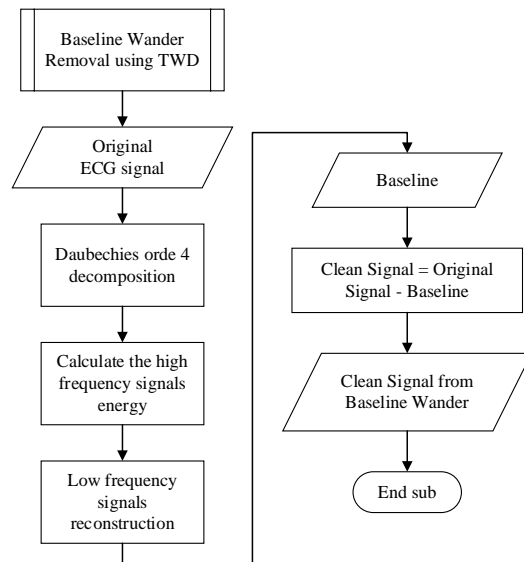


Figure 5. Baseline Wander Removal Flowchart

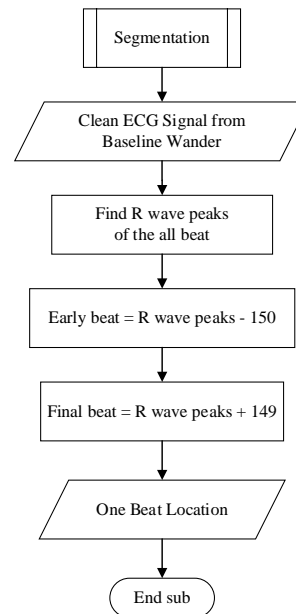


Figure 6. Beat Segmentation Flowchart

ssification results. Figure 4 shows the testing system flowchart.

### Baseline Wander Removal

Baseline Wander is a condition when the ECG signal is shifted up or down to the isoelectric line (line axis). To overcome the noise, the researchers conducted the Baseline Wander Removal using Transform Wavelet Discrete method [6].

The original signal is decomposed using the Transform Wavelet type Daubechies order 4. The system calculated the energy value at a high-frequency signal and found the conditions in which

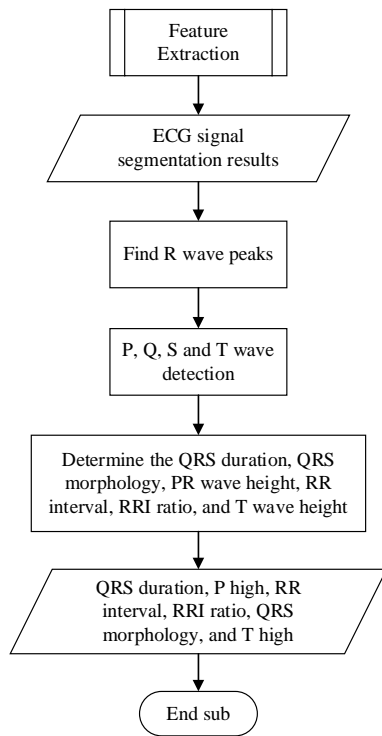


Figure 7. Feature Extraction Flowchart

the value of decomposition level is lower than the value of decomposition level before. Furthermore, the system performed approximation signal reconstruction of this level by removing the value of the high-frequency signal. Figure 5 shows the Baseline Wander Removal flowchart.

**Beat Segmentation**

In this stage, the ECG signals segmented into each beat. This process used additional annotation file to locate the position of the R-wave peaks. Assuming the width of each beat is done by positioning the peak of R as a pivot for each beat [8]. Early signals are sliced starting from the position of R-150 to R-149, so we get as many as 300 samples of beat data. Figure 6 shows the beat segmentation flowchart.

**Feature Extraction**

Feature extraction is the process of detection characteristics of the electrocardiogram that is used as an input variable in the classification process. The process consists of several steps as shown in Figure 7. This process begins by determining the R-wave peaks on the signal segmentation results, and then continued for the wave that is P, Q, S and T. After all waves found, then continued to characterize the one beat of ECG signal. These include the QRS duration, QRS morphology, P

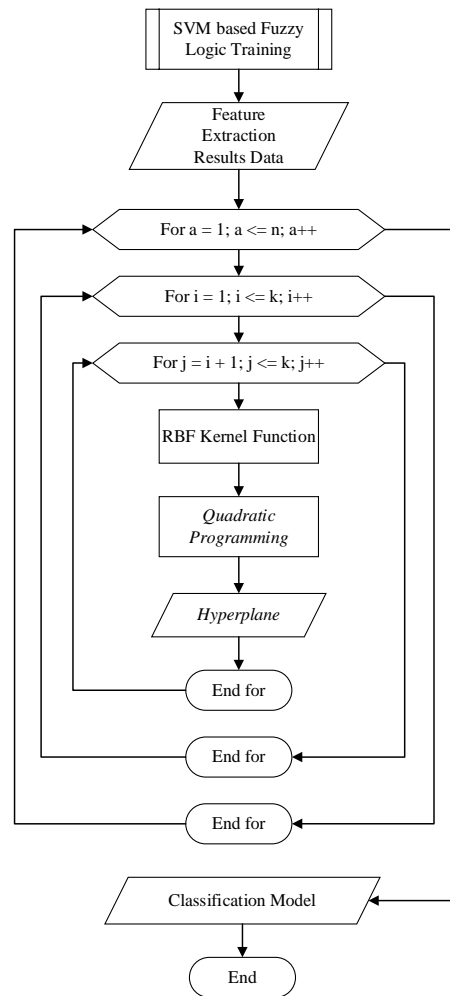


Figure 8. Training Process Flowchart

wave height, RR interval, RRI ratio, and T wave height.

**Training Process**

SVM based fuzzy logic training process used the same training process with SVM one against one method, and build  $k(k - 1)/2$  parts of the binary classification model (k is the classes number). Each classification models is trained using data from two classes. For classroom training data to class i and class j, the system will solve the problems with the 6 classes by finding 15 separator functions as follows:

$$D_{ij}(x) = D_{12}(x), D_{13}(x), D_{14}(x), D_{15}(x), D_{16}(x), D_{23}(x), D_{24}(x), D_{25}(x), D_{26}(x), D_{34}(x), D_{35}(x), D_{36}(x), D_{45}(x), D_{46}(x), D_{56}(x).$$

So that the data can be separated linearly, this system uses Radial Basis Function (RBF) kernel  $K(x, y) = \exp(\frac{-1}{2\sigma^2} \|x - y\|^2)$ . To obtain a solution or the separation of these two classes, this system using quadratic programming. Result of

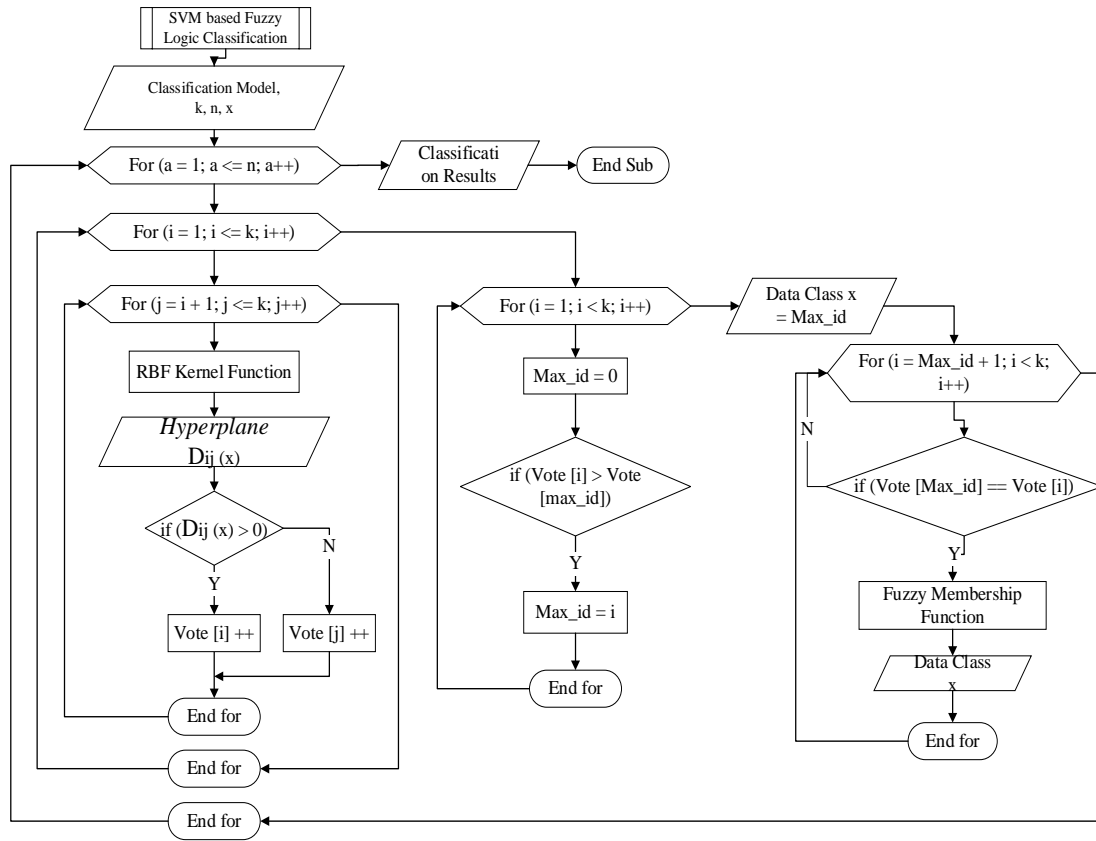


Figure 9. Classification Process Using SVM based fuzzy logic

quadratic programming are  $w$ ,  $x$ , and  $B$  that used for the testing process. The flowchart for the stages of the training process can be seen in Figure 8.

### Classification Process Using SVM based fuzzy logic

A common strategy in the testing process using one-against-one SVM method is max-voting [9]. Based on this strategy, if the representation of a data on the hyperplane  $D_{ij}(x)$  is in class  $i$ , then voting added one for the class  $i$ . Otherwise, if the representation of a data on the hyperplane  $D_{ij}(x)$  is in class  $j$ , then voting added one for the class  $j$ . These steps are repeated for all hyperplane. Then, system predicted  $x$  is in a class where, based on the value of the highest voting. In cases where there are two classes with the same voting value, then selected the smallest index value, the process is arguably less effective. To resolve with the data that is not classifiable, this study used the membership function according to the equation(11) using  $m_i(x)$ . The data  $x$  are classified based on the highest membership value according to the equation(12). Flowchart for the stages SVM based fuzzy logic method can be seen in Figure 9.

### 3. Results and Analysis

Evaluations of the classification results of the arrhythmia abnormalities consisted of two scenarios, among others, using the original SVM, and proposed method (SVM based fuzzy logic). This study used 600 testing data, which consisted of each 100 Normal Beat (NB) data, 100 Left Bundle Branch Block Beat (LBBB) data, 100 Right Bundle Branch Block Beat (RBBB) data, 100 Premature Ventricular Contraction (PVC) data, 100 Atrial Premature Beat (APB) data, and 100 Paced Beat (PB) data. Then the results are compared based on three measures among others accuracy, sensitivity, and specificity.

$$\text{Accuracy} = \frac{\text{number of data classified correctly}}{\text{number of tests performed}} \times 100\% \quad (13)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \quad (14)$$

$$\text{Specificity} = \frac{TN}{FP + TN} \times 100\% \quad (15)$$

where True Positive (TP) is arrhythmia correctly

TABLE 1  
CLASSIFICATION RESULTS USING SVM METHOD

	Arrhythmia Type / Class					
	NB	LBBB	RBBB	PVC	PB	APB
TP	93	95	87	91	91	94
FN	7	5	13	9	9	6
FP	14	6	5	7	10	7
TN	486	494	495	493	490	493

identified, False Positive (FP) is arrhythmia wrongly identified, True Negative (TN) is arrhythmia correctly not identified, and False Negative (FN) is arrhythmia wrongly not identified.

#### Scenario 1: Classification using SVM Method [4]

Experimental results of classification using SVM method [4] can be seen in Table 1.

$$\text{Sensitivity of NB Class} = \frac{93}{93+7} \times 100\% = 93\%$$

$$\text{Sensitivity of LBBB Class} = \frac{95}{95+5} \times 100\% = 95\%$$

$$\text{Sensitivity of RBBB Class} = \frac{87}{87+13} \times 100\% = 87\%$$

$$\text{Sensitivity of PVC Class} = \frac{91}{91+9} \times 100\% = 91\%$$

$$\text{Sensitivity of PB Class} = \frac{91}{91+9} \times 100\% = 91\%$$

$$\text{Sensitivity of APB Class} = \frac{94}{94+6} \times 100\% = 94\%$$

Sensitivity values for each class are summed and divided by six, to obtain the average sensitivity value of 91.83%.

$$\text{Specificity of NB Class} = \frac{486}{14+486} \times 100\% = 97.2\%$$

$$\text{Specificity of LBBB Class} = \frac{494}{6+494} \times 100\% = 98.8\%$$

$$\text{Specificity of RBBB Class} = \frac{495}{5+495} \times 100\% = 99\%$$

$$\text{Specificity of PVC Class} = \frac{493}{7+493} \times 100\% = 98.6\%$$

$$\text{Specificity of PB Class} = \frac{490}{10+490} \times 100\% = 98\%$$

$$\text{Specificity of APB Class} = \frac{493}{7+493} \times 100\% = 98.6\%$$

The classification system using SVM method has average specificity value of 98.36%. From Table 1, we can calculate the overall accuracy value of 91.83%, the sensitivity average value of 91.83%, and the average specificity value of 98.36%.

#### Scenario 2: Classification using Proposed Method

Experimental results of classification using proposed method (Support Vector Machine based fuzzy logic method) can be seen in Table 2. Experiments scenario 2 uses the same dataset as that used in scenario 1.

TABLE 2  
CLASSIFICATION RESULTS USING PROPOSED METHOD

	Arrhythmia Type / Class					
	NB	LBBB	RBBB	PVC	PB	APB
TP	95	95	93	91	93	94
FN	5	5	7	9	7	6
FP	10	6	5	7	4	7
TN	490	494	495	493	496	493

$$\text{Sensitivity of NB Class} = \frac{95}{95+5} \times 100\% = 95\%$$

$$\text{Sensitivity of LBBB Class} = \frac{95}{95+5} \times 100\% = 95\%$$

$$\text{Sensitivity of RBBB Class} = \frac{93}{93+7} \times 100\% = 93\%$$

$$\text{Sensitivity of PVC Class} = \frac{91}{91+9} \times 100\% = 91\%$$

$$\text{Sensitivity of PB Class} = \frac{93}{93+7} \times 100\% = 93\%$$

$$\text{Sensitivity of APB Class} = \frac{94}{96+4} \times 100\% = 94\%$$

Sensitivity values of classification using proposed method for each class are summed and divided by six, to obtain the average sensitivity value of 93.5%.

$$\text{Specificity of NB Class} = \frac{490}{10+490} \times 100\% = 98\%$$

$$\text{Specificity of LBBB Class} = \frac{494}{6+494} \times 100\% = 98.8\%$$

$$\text{Specificity of RBBB Class} = \frac{495}{5+495} \times 100\% = 99\%$$

$$\text{Specificity of PVC Class} = \frac{493}{7+493} \times 100\% = 98.6\%$$

$$\text{Specificity of PB Class} = \frac{496}{4+496} \times 100\% = 99.2\%$$

$$\text{Specificity of APB Class} = \frac{493}{7+493} \times 100\% = 98.6\%$$

The classification system of classification using proposed method has average specificity value of 98.7%. From Table 2, we can calculate the overall accuracy value of 93.5%, the average sensitivity value of 93.5%, and the average specificity value of 98.7%. These results indicate that the arrhythmia classification system uses proposed method deliver performance values of accuracy, sensitivity, and specificity higher than the arrhythmia classification system using the original SVM method. This is because the use of fuzzy membership functions can cope with the data that is not classifiable in one against one Support Vector Machine method.

#### 4. Conclusion

From the results of experiments and calculations, it can be concluded that the use of support vector machine based on fuzzy logic methods for classification system abnormalities on ECG arrhythmia delivered performance with an accuracy level of 93.5%, the sensitivity of 93.5% and specificity of 98.7%. The result is higher when compared with the use of Support Vector Machine method with

an accuracy of 91.83%, sensitivity of 91.83%, and specificity of 98.36%. The use of fuzzy membership function can cope with data that are not classifiable in the one against one SVM method to deliver performance values of accuracy, sensitivity, and specificity higher than SVM method. This research is only using one lead of ECG signals. Some arrhythmia disorders common feature but is different leads, such as abnormalities in Left Bundle Branch Block Beat (LBBB), and Beat Right Bundle Branch Block (RBBB). Further research can be done using two leads to improve the accuracy of the arrhythmia classification system.

## References

- [1] A. Dallali, A. Kachouri, & M. Samet. "Fuzzy c-means clustering, Neural Network, WT, and HRV for classification of cardiac arrhythmia", *ARNP Journal of Engineering and Applied Sciences*. Vol. 6(10). 2011.
- [2] Arrhythmia Rules. Biomedical Telemetry. <http://www.biotel.ws/ekgs/TheRules.htm> accessed on 12 May 2014.
- [3] P. Srivastava, N. Sharma, & C.S. Aparna. *Fuzzy Soft System and Arrhythmia Classification*. Hindawi. India. 2013.
- [4] M. Moavenian, & H. Khorrami. "A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification". Elsevier. 3088-3093. 2010.
- [5] O. Castilo, P. Melin, E. Ramirez, & J. Soria. "Hybrid intelligent system for cardiac arrhythmia classification with Fuzzy K-Nearest Neighbours and neural networks combined with a fuzzy system". Elsevier. 2947-2955. 2012.
- [6] D. Sripathi. *Efficient Implementations of Discrete Wavelet Transform Using FPGAs*. Florida State University. Florida. 2003.
- [7] A.K.M.F. Haque, H. Ali, M. A. Kiber, & Md. T. Hasan. "Detection of Small Variations of ECG Features Using Wavelet", *ARNP Journal of Engineering and Applied Sciences*. 4(6):27-30. 2009.
- [8] B. Anuradha, & V.C. Veera Reddy. "ANN for Classification of Cardiac Arrhythmias", *ARNP Journal of Engineering and Applied Sciences*. 3(3):1-6. 2008.
- [9] A. Dallali, A. Kachouri, & M. Samet. "Integration Of HRV, WT and Neural Networks for ECG Arrhythmia Classification", *ARNP Journal of Engineering and Applied Sciences*. 06/2011; 6(5). 2011.
- [10] "MIT-BIH Arrhythmia database," 2015. [Online]. Available:<http://www.physionet.org/physiobank/database/mitdb/>. Accessed on 29 May 2014.
- [11] K. J. Reza, S. Khatun, M. F. Jamlos, M. M. Fakir, & S. S. Mostafa. "Performance evaluation of diversified SVM kernel functions for breast tumour early prognosis", *ARNP Journal of Engineering and Applied Sciences*. 03/2014; 9(3):329-335. 2014.
- [12] T. Kitamura, S. Abe, & K. Fukui. *Subspace Based Least Squares Support Vector Machines For Pattern Classification*. IEEE International Joint Conference on Neural Networks, 2009.