Weighted SVM with RR Interval based Features for Android-based Arrhythmia Classifier

Muhammad Ilham Rizqyawan, Artha Ivonita Simbolon, and Dwi Esti Kusumandari
Technical Implementation Unit for Instrumentation Development
Indonesian Institute of Sciences
Bandung, Indonesia

***Abstract*— ECG is one of the most popular fields in biosignals research. One of the popular area in ECG research is automatic Arrhythmia classification. In this paper, we presented an effort to make an Arrhythmia classifier for Android. We use RRI based features and SVM as the classification method. Then we conduct an experiment with three different SVM configuration to see how much improvement can be made by using these configurations. By looking at kappa score as the metrics, the configuration 2 is greatly improve the classifier (169% increase). And by using hyper-parameter tuning we further optimize the classifier as can be seen on result of configuration 3 (10.5% increase).**

***Index Terms*—Android, Arrhythmia, ECG, SVM**

I. Introduction

BIOSIGNAL is a signal generated by living beings where the signal can be measured and monitored to see its health level. One of the most popular in electrical biosignal to be measured and researched is the Electrocardiogram (ECG) [1]. ECG is a non-invasive tool for detecting the electrical activity of the heart [2]–[5]. Research on ECG is one of the important studies, because it can detect cardiovascular diseases[1]. Cardiovascular diseases can be characterized by arrhythmia.

Arrhythmias can occur suddenly and anywhere. Often many patients are late for the detection of arrhythmias, and even people die suddenly without treatment. So it takes a tool that can detect the arrhythmia quickly and send it to the nearest doctor. Therefore, there has been much research on the development of ECG instruments, both in terms of hardware [6], [7] and software [8]–[11].

On the other hand, as technology develops, there are many Android based health applications[12], [13]. Therefore, we are creating an Android-based ECG based app. With this application, can be done with the prevention of cardiovascular diseases with the detection of arrhythmias. Inside the ECG application, there is an ECG signal processing so the display on the application is easily understood by anyone.

In the processing of ECG signals, the features extraction and classification are the important steps. There have been many types of feature extraction used in ECG research, such as wavelet transform coefficient features[14]–[17], frequency-based features [18], [19] and Hermite polynomials [20], [21]. Most of them use either time- or frequency- domain representation of the ECG signals as features. Depending on the feature, classification is allowed to recognize the class. Classification is commonly used in ECG research is neural network (NN) [22], support vector machine (SVM)[2], [23], fuzzy clustering neural network (FCNN) [24], and artificial neural networks (ANN) [25].

In this paper, our focus is on developing ECG signals for Android. We use RRI for feature and SVM as a classification method. An experiment was conducted with three different SVM configurations to see how many improvements can be made using this configuration.

II. Methodology

This research project consists of two sub-research focuses, which is signal processing and system development. The system development stage is discussed in our other paper [1]. Signal processing stage consisting of all the steps required to recognize the pattern that occurs in the ECG signal. In this stage, we investigated which method is effective and efficient in each of the steps in order to classify the ECG signal.

As we can see in Fig. 1., arrhythmia classifier for the android application is developed separately in the desktop workstation. This is important since we need a lot of computation power to train the classifier model. We use MIT-BIH - which is one of the standard database for Arrhythmia - as training data. After the classifier is trained, the model is exported to be applied in the android application.

This paper will be focused on the classifier development part, specifically the SVM. We focus to develop classifier with minimal feature to make it as light as possible in the computation as the classifier will be applied in the android application. In this paper, we compare the performance of 3 different configurations of SVM.

Manuscript received August 15, 2017. This research was supported by the Insinas Ristek funded by Minister of Research, Technology, and Higher Education, Indonesia.

Muhammad Ilham Rizqyawan, Artha Ivonitha Simbolon, and Dwi Esti Kusumandari is with the Technical Implementation Unit for Instrumentation Development, Indonesian Institute of Sciences, Indonesia. (corresponding author to provide phone: +6222-2503053; e-mail: muha176@lipi.go.id).

ISSN: 1942-9703 / CC BY-NC-ND

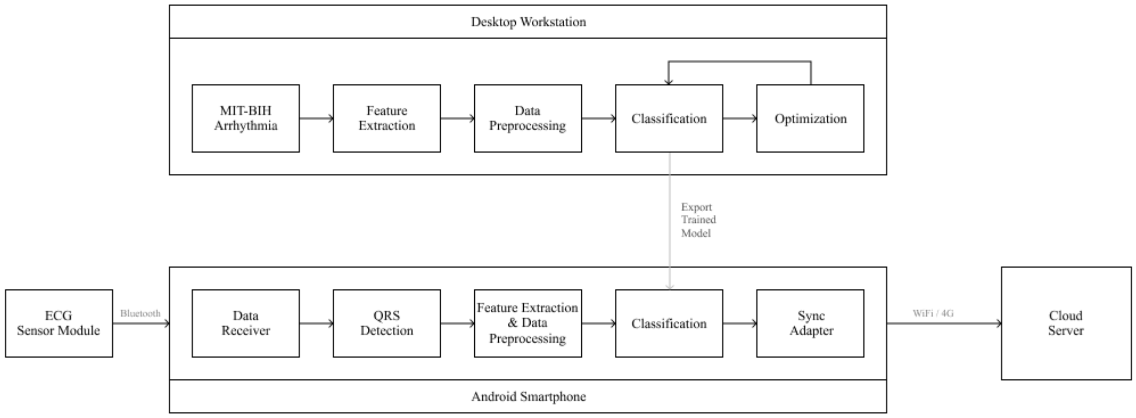

Fig. 1. Big picture of the research project. The classifier is developed in desktop workstation to be exported later to the android application.

A. MIT-BIH Database

MIT-BIH [26] is one of the standard databases recommended by the AAMI for the standard Arrhythmia evaluation, therefore most of the published research on Arrhythmia used it for their database. MIT-BIH contains 48 records, with each record length is approximately 30 minutes long. Each record contains two channel ambulatory ECG recordings. The sampling rate for the recordings is 360Hz with 11-bit resolution over a 10mV range. Each beat in the records is annotated by two or more cardiologists independently. The total of annotated beats of the 48 records is approximately 110.000 annotations. This database is obtained from Physionet [27], [28].

In AAMI recommended practice report [29], AAMI recommends the classes for the arrhythmia classification: Normal (N), Supraventricular ectopic beat (SVEB), Ventricular ectopic beat (VEB), Fusion beat (F), and Unknown beat (Q). Details on this class can be seen in Table I. In this study, we only focused to make classifier for the first three class (N, SVEB, and VEB), although we still include the F and Q class in the classifier. We prioritize N, SVEB, and VEB classes as the number of data on the last two class is very limited.

AAMI also recommends removing four of the recordings (102, 104, 107, and 217) from the analysis since it contains paced beats. By removing these recording, the total of beats in training data is 100.737 beats. As we can see in Fig. 2 this is a very imbalance data, with almost 90% of its beat is normal class. With this kind of training data, the classifier will likely be biased to the majority class. Our effort to minimize the bias will be discussed in the next section.

TABLE I
MIT-BIH ARRHYTHMIA DATABASE GROUPED BY AAMI CLASSES

Class	Description	Physionet's Annotation	MIT-BIH Types
N	Normal	N	Normal beat (NOR)
		L	Left bundle branch block beat (LBBB)
		R	Right bundle branch block beat (RBBB)
		e	Artial escape beat (AE)
		j	Nodal (junctional) escape beat (NE)
SVEB	Supraventricular ectopic beat	A	Artial premature beat (AP)
		a	Aberrated atrial premature beat (aAP)
		J	Nodal (junctional) premature beat (NP)
		S	Supraventricular premature beat (SP)
VEB	Ventricular ectopic beat	V	Premature ventricular contraction (PVC)
		E	Ventricular escape beat (VE)
F	Fusion beat	F	Fusion of ventricular and normal beat (fVN)
Q	Unknown beat	/	Paced beat (P)
		f	Fusion of paced and normal beat (fPN)
		Q	Unclassified beat (U)

ISSN: 1942-9703 / CC BY-NC-ND

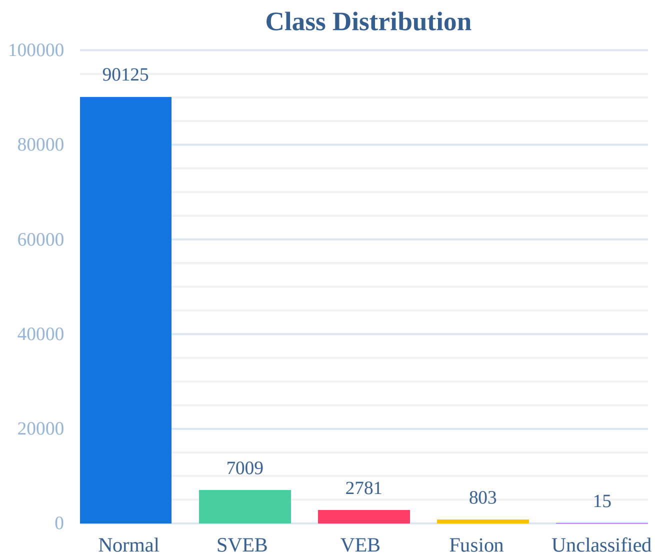

Fig. 2. Class distribution of training data. As we can see the normal class is dominated with almost 90% of the data.

B. Feature Extraction

In this research, we took five features that consist of peak amplitude, local average, R-to-R interval (RRI), R-to-R interval ratio (RRIR), and 10RRIR. Peak amplitude is value of each R-peak. Local average is the average of ten of interval between R-peak. First, select R-peak, then calculate five previous R-peak intervals and five R-peak intervals next to it. The equation of local average is showed in Equation (1), where LA is local average, x is R-peak, and i is the R-peak position.

$$LA = \frac{1}{10}\sum_{i-5}^{i+5} x_i - x_{i-1} \tag{1}$$

RRI is divided into six that is pre-RR, pre-RR2, pre-RR3, post-RR, post-RR2 and post-RR3. Pre-RR is the interval between selected R-peak and R-peak before. Pre-RR2 is the interval between selected R-peak and two R-peak before. Pre-RR3 is the interval between selected R-peak and three R-peak before. Post-RR is the interval between selected R-peak and R-peak next. Post-RR2 is the interval between selected R-peak and two R-peak next. Post-RR3 is the interval between selected R-peak and three R-peak next. Illustration for RRI is showed in Fig. 3.

RRIR is the ratio between the pre-RR with pre-RR next (see Equation (2)). 10RRIR is ratio between the pre-RR with interval of the R-peak to ten R-peak next (see Equation (3)).

$$RRIR = \frac{pre{-}RR_i}{pre{-}RR_{i+1}} \tag{2}$$

$$10RRIR = \frac{pre{-}RR_i}{(|x_i - x_{i+10}|)} \tag{3}$$

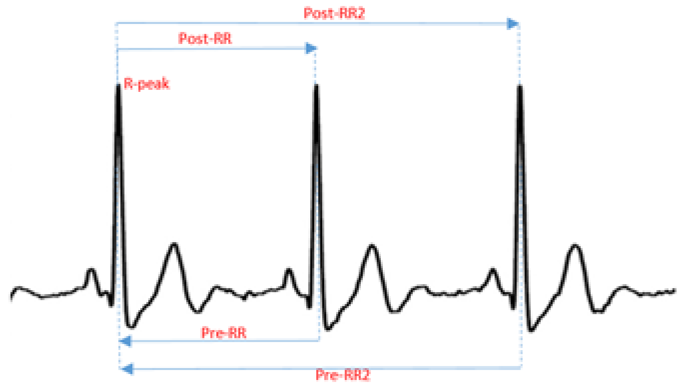

Fig. 3. Illustration of Extraction Feature

RRI is a feature selected from all the above features. Because this research aims to obtain a simple process and not burdensome Android.

C. Data Preprocessing

The results of feature extraction are normalized in order to reduce complexity. Equation (4) is the normalization equation used

$$z_i = \frac{x_i - min(x)}{max(x) - min(x)} \tag{4}$$

D. Classification

SVM is the classification methodology introduced by Vapnik [30], [31]. In a simple, the SVM identifies the best separating hyperplane between the two classes of the training samples with the plane used is the maximum margins. In this way, not only an optimal hyperplane is fitted, but also less training samples are effectively used. Thus, high classification accuracy can be achieved with small training sets. Given a set of training (x_i, y_i), $i = 1,2,\dots,l$, where $x_i \in R^n$ and $y_i \in \{-1,1\}$, the SVM algorithm is summarized as the following optimization problem:

$$\underset{w,b,\xi}{min} \left\{\frac{1}{2}\|w\|_2^2 + C\left(\sum_{i=1}^{l}\xi_i\right)\right\} \tag{5}$$

constrained to

$$y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i,\ \forall i = 1,\dots,n$$
$$\xi_i \geq 0,\ \forall i = 1,\dots,n$$

where $w, b,$ and ξ_i are the weight vector, bias and slack variable, respectively. C is a constant and determined a priori. $\phi(x)$ is a nonlinear function that maps x into a higher dimensional space.

Searching for the optimal hyperplane in Equation (5) is a quadratic programming problem, which can be solved by constructing a Lagrangian and transforming it into a dual maximization problem of the function $Q(a)$, defined as follows:

$$\max Q(a) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (6)$$

constrained to

$\sum_{i=1}^{l} \alpha_i y_i = 0$;

$0 \le \alpha_i \le C$, for $i = 1,2, \dots, l$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)$ is the vector of non-negative Lagrange multipliers. And the Kernel function is determined as follows

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (7)$$

Assuming that the optimum values of the Lagrange multipliers are denoted as $\alpha_{o,i} (i = 1,2, \dots, l)$, then possible to determine the fit optimum value of the optimum linear w_o and hyperplane vectors as in (8) and (9), respectively:

$$w_o = \sum_{i=1}^{l} \alpha_{o,i} y_i \phi(x_i) \quad (8)$$

$$\sum_{i=1}^{l} \alpha_{o,i} y_i K(x, x_i) + b = 0 \quad (9)$$

The decision function can be written as

$$f(x) = sgn\left(\sum_{i=1}^{l} \alpha_{o,i} y_i K(x, x_i) + b\right) \quad (10)$$

In this research, the radial basis function (RBF) is used as the kernel function and the parameters-kernel width σ and regularization constant C were experimentally defined to achieve the best classification result. RBF is showed in Equation (11) [32].

$$K(x_i, x_j) = exp\left(-\frac{\|x_i - x_j\|}{2\sigma^2}\right), \ \sigma \in \mathbb{R} \quad (11)$$

Then did the optimization by using Gaussian Process. Gaussian Process (GP) is a function that makes it easy and great in the problem of distribution function. GP can be used to describe a distribution function.

$$f \sim GP(m, K) \quad (12)$$

where m: X → R is a mean function.

$$m(x) = E\,[f(x)] \quad (13)$$

and K: X2 → R is a covariance function

$$K(x, x') = E\,[(f(x) - m(x)(f(x') - m(x'))] \quad (14)$$

GP performed as many as 30 iterations with Cohen's Kappa as the constanta function to obtain the C and gamma values as follows (conf 3).

TABLE II
SVM CLASSIFIER CONFIGURATIONS

Title	C	Gamma	Weight
Configuration 1	1.0	0.0	1:1:1:1:1
Configuration 2	1.0	0.0	1:13:33:1:1
Configuration 3	9.5	9.5	1:13:33:1:1

III. RESULT

Using SVM with default parameter (C: 1.0, gamma: 0.0 and no class weighting) we got a relatively good accuracy score of 89.85%. The details of the score can be seen in Table III.

TABLE III
ACCURACY OF SVM CONFIGURATION 1

Correctly classified instance	90507	89.8484%
Incorrectly classified instance	10226	10.1516%

But looking at the accuracy as the sole metric for the classifier evaluation is not enough and can be misleading, especially on imbalance data. So, we use other metrics recommended by AAMI for evaluating methods: Sensitivity (Se), Positive predictivity (+P), and False positive rate (FPR) in addition to overall accuracy

If True Positive (TP) is number of correctly identified instances, False Positive (FP) is incorrectly identified instances, True Negative (TN) is correctly rejected instances, and False Negative (FN) is incorrectly rejected instances; these metrics can be calculated from confusion matrix as following:

$$Se = \frac{TP}{(TP+FN)} \quad (15)$$

$$+P = \frac{TP}{(TP+FP)} \quad (16)$$

$$FPR = \frac{FP}{(FP+TN)} \quad (17)$$

We also use other metrics to further evaluate the classifier, which is very handy to evaluate classifier with imbalance data: Cohen's Kappa statistics [33], and Area under Receiver Operating Characteristic (AUROC) [34]. Kappa can be calculated by following:

$$Kappa = \frac{p_o - p_e}{1 - p_e} \quad (18)$$

$$p_o = \frac{TP+TN}{TP+TN+FP+FN} \quad (19)$$

$$p_e = \frac{(TN+FP)*(TN+FN)+(FN+TP)*(FP+TP)}{Total*Total} \quad (20)$$

The results for SVM configuration 1 are shown in confusion matrix in Table III. And for the analysis, the results in the form of all mentioned measurement metrics above can are shown in Table IV.

TABLE IV
RESULTS FOR EACH SVM CONFIGURATION WITH OTHER METRICS

		Configuration 1	Configuration 2	Configuration 3
Acc (%)		89.8484	71.0274	73.00
Kappa		0.0863	0.2322	0.26
Se (%)	N	99.84	72.23	74.26
	SVEB	0.00	68.51	70.62
	VEB	18.99	59.30	59.55
	F	0.00	0.00	0.00
	Q	0.00	0.00	0.00
+P (%)	N	89.93	96.07	96.31
	SVEB	~	17.04	18.61
	VEB	78.34	34.39	35.64
	F	~	~	~
	Q	~	~	~
FPR (%)	N	95.02	25.13	24.14
	SVEB	0.00	24.94	23.09
	VEB	0.15	3.21	3.05
	F	0.00	0.00	0.00
	Q	0.00	0.00	0.00
AUROC	N	0.52	0.74	0.75
	SVEB	0.50	0.72	0.74
	VEB	0.59	0.78	0.78
	F	0.50	0.50	0.50
	Q	0.50	0.50	0.50

(~) symbol in the positive predictivity data (+P) is caused by zero count in both of TP and FP.

From table IV. we can see that accuracy of configuration 1 result is misleading. If we look at the breakdown of each class in confusion matrix, we can see that the classifier mistook 95% of beats as N class. This can be dangerous especially in the medical application. In this area of application, False Positive in the normal / healthy class needs to be kept at minimum, since this could lead to omission of the patient who actually need a treatment.

One of the major suspect of this poor classifier bias is the imbalance class problem we mentioned earlier. If we look at the configuration 2, by using weighting on the SVM, the FPR of the N class is reduced and the sensitivity of the SVEB and VEB class is improved significantly. Another way to see the improvements of the classifier without having to check the metrics of each class is by using kappa statistics. As we can see, the kappa is also improved in the configuration 2.

In the configuration 3, we further optimize the classifier by tuning the hyper-parameter. As we can see the classifier is improved which can be seen in the kappa and AUROC score, although the improvements are not significant. We suspect that this is due to limited searching range and number of iteration of the Gaussian Process. So, we suggest using a boarder searching range and use more iteration on the optimization, although this also means spending more time in the hyper-parameter tuning.

IV. CONCLUSION

In this paper, we presented an effort to make an Arrhythmia classifier for Android by using RR based features and three configurations of SVM. The overall accuracy of configuration 1 is higher than the other 2, but it is misleading by only seeing this metric. By looking at other metrics such as kappa and AUROC score, the configuration 2 is greatly improve the classifier (169% increase on Kappa score). And by using hyper-parameter tuning we further optimize the classifier as can be seen on result of configuration 3 (10.5% increase on Kappa score).

While the overall performance of the classifier is not good enough, but the use of weighted SVM and Gaussian Process for hyper-parameter tuning is quite promising to be considered in future work. We suggest trying different features to improve the overall performance of the classifier, by still considering the performance issue in mind, since the aim of this study is to develop classifier for mobile device. Other technique such as undersampling and oversampling (i.e. SMOTE [35]) can be used to improve the classifier with imbalance data. And finally, to make sure the classifier is well tested, we also suggest using an inter-patient scheme by dividing the training and testing data, such as protocol proposed by [9]. This protocol is more realistic to see the actual performance of the classifier since the classifier could be overfitting to the training data.

ACKNOWLEDGMENT

Authors wish to thank Bandung Technical Management Unit for Instrumentation Development (Deputy for Scientific Services) for supporting in facility and equipment.

REFERENCES

[1] M. I. Rizqyawan, M. F. Amri, R. P. Pratama, and A. Turnip, “Design and development of Android-based cloud ECG monitoring system,” *Proc. - 2016 3rd Int. Conf. Inf. Technol. Comput. Electr. Eng. ICITACEE 2016*, pp. 1–5, 2017.

[2] M. K. Das and S. Ari, “Patient-specific ECG beat classification technique,” *Healthc. Technol. Lett.*, vol. 1, no. 3, pp. 98–103, 2014.

[3] Z. H. Slimane and A. Naït-ali, “QRS complex detection using Empirical Mode Decomposition,” *Digit. Signal Process.*, vol. 20, no. 4, pp. 1221–1228, 2010.

[4] R. Alonso, A. J. Méndez, and X. A. Vila, “A comparison of three QRS detection algorithms over a public database,” *Procedia Technol.*, vol. 9, pp. 1159–1165, 2013.

[5] Z. Zhang, J. Dong, X. Luo, K. Choi, and X. Wu, “Heartbeat

classification using diseasespecific feature selection," vol. 46, pp. 79–89, 2014.
[6] S. Das; R. Gupta; M. Mitra "Development of Analog ECG simulator using standalone embedded system," *International Journal of Electrical, Electronics and Computer Engineering*, vol. 1 no. 2, pp. 83-87, 2017.
[7] S. Borromeo, C. Rodriguez-Sanchez, F. Machado, J. A. Hernandez-Tamames, and R. De La Prieta, "A reconfigurable, wearable, wireless ECG system," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, pp. 1659–1662, 2007.
[8] D. Hayn, B. Jammerbund, and G. Schreier, "QRS detection based ECG quality assessment," *Physiol. Meas.*, vol. 33, no. 9, pp. 1449–1461, 2012.
[9] P. De Chazal, M. O. Dwyer, R. B. Reilly, and S. Member, "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features," vol. 51, no. 7, pp. 1196–1206, 2004.
[10] S. Pal and M. Mitra, "Empirical mode decomposition based ECG enhancement and QRS detection," *Comput. Biol. Med.*, vol. 42, no. 1, pp. 83–92, 2012.
[11] S. Gradl, P. Kugler, C. Lohmuller, and B. Eskofier, "Real-time ECG monitoring and arrhythmia detection using Android-based mobile devices," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 2452–2455, 2012.
[12] M. Altini, J. Penders, and H. Roebbers, "An Android-based body area network gateway for mobile health applications," *Wirel. Heal. 2010 - WH '10*, p. 188, 2010.
[13] C. Liu, Q. Zhu, K. A. Holroyd, and E. K. Seng, "Status and trends of mobile-health applications for iOS devices: A developer's perspective," *J. Syst. Softw.*, vol. 84, no. 11, pp. 2022–2033, 2011.
[14] P. Li, M. Liu, X. Zhang, and H. Da Chen, "A low-complexity ECG processing algorithm based on the Haar wavelet transform for portable health-care devices," *Sci. China Inf. Sci.*, vol. 57, no. 12, pp. 1–14, 2014.
[15] M. Go and F. Ata, "Denoising of weak ECG signals by using wavelet analysis and fuzzy thresholding," no. 2000, 2012.
[16] Z. Zidelmal, A. Amirou, M. Adnane, and A. Belouchrani, "QRS detection based on wavelet coefficients," *Comput. Methods Programs Biomed.*, vol. 107, no. 3, pp. 490–496, 2012.
[17] S. Sumathi, H. L. Beaulah, and R. Vanithamani, "A Wavelet Transform Based Feature Extraction and Classification of Cardiac Disorder," 2014.
[18] R. Vijay, "Time-Frequency Filtering with the S-Transform of ECG Signals," vol. 3, no. 2, pp. 1–5, 2013.
[19] K. I. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with fourier-transform neural network," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 179–185, 1999.
[20] M. Lagerholm and G. Peterson, "Clustering ECG complexes using hermite functions and self-organizing maps," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 838–848, 2000.
[21] K. S. Park, B. H. Cho, D. H. Lee, S. H. Song, J. S. Lee, Y. J. Chee, I. Y. Kim, and S. I. Kim, "Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function," *2008 Comput. Cardiol.*, pp. 229–232, 2008.
[22] T. Lahiri, S. Sarkar, S. Sanyal, A. A. Morozov, and Y. V Obukhov, "Clustering of Signal Components within Most Likely ECG Episodes to Analyze the ECG-Waves 1," vol. 19, no. 1, pp. 30–34, 2009.
[23] C. Engineering, "Electrocardiogram beat classification using s-transform based feature set," vol. 14, no. 5, pp. 1–18, 2014.
[24] Y. Özbay, R. Ceylan, and B. Karlik, "A fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Comput. Biol. Med.*, vol. 36, no. 4, pp. 376–388, 2006.
[25] G. K. Prasad and J. S. Sahambi, "ECG Arrhythmias using Multi-Resolution Analysis and Neural Networks," *TENCON 2003. Conf. Converg. Technol. Asia-Pacific Reg.*, no. 3, pp. 227–231, 2003.
[26] G. B. Moody and R. G. Mark, "The Impact of the MIT-BIH Arrhythmia Database," no. June, pp. 45–50, 2001.
[27] S. H. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, "PhysioBank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.
[28] "PhysioNet the research resource for complex physiologic signals." [Online]. Available: https://www.physionet.org.
[29] American National Standards Institute, "Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms: ANSI/AAMI EC57," *Assoc. Adv. Med. …*, 2012.
[30] G. De Lannoy, D. Franc, J. Delbeke, and M. Verleysen, "Weighted SVMs and Feature Relevance Assessment in Supervised Heart Beat Classification," pp. 212–223, 2011.
[31] M. Moavenian and H. Khorrami, "A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification," *Expert Syst. Appl.*, vol. 37, no. 4, pp. 3088–3093, 2010.
[32] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, 2005.
[33] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, pp. 37–46, 1960.
[34] A. J. Hanley and J. B. McNeil, "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve," *Radiology*, vol. 143, pp. 29–36, 1982.
[35] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

Muhammad Ilham Rizqyawan received the B.Comp degrees in computer sciences / informatics from Indonesian Computer University (UNIKOM) in 2013 and M.Eng. in electrical engineering from Institut Teknologi Bandung (ITB) in 2017. He currently works in the Technical Implementation Unit for Instrumentation Development, Indonesian Institute of Sciences, Indonesia. His research interest includes human-computer interaction, data visualization, game-based learning, internet of things, and machine learning.

Artha Ivonita Simbolon received the B.Sc degree in Physics from Institute Technology of Bandung (ITB). She is currently work in the Technical Implementation Unit for Instrumentation Development, Indonesian Institute of Sciences, Indonesia as junior researcher. Her research interests are in electronics and control engineering, artificial intelligent, signal processing, and brain-computer interface.

Dwi Esti Kusumandari received the B.Eng. degrees in Engineering Physics from the Institute of Technology Bandung (ITB), Indonesia, in 1999, and the M.Eng. degree in Biomedical Engineering from the same institute in 2006. She is currently work in the Technical Implementation Unit for Instrumentation Development, Indonesian Institute of Sciences, Indonesia as a junior researcher. Her research interests are in biomedical engineering, image processing, signal processing, and instrumentation.