

# A Method of Analyzing ECG to Diagnose Heart Abnormality utilizing SVM and DWT

Ahmed Shdefat<sup>1</sup>, Moonil Joo<sup>1</sup>, Heecheol Kim<sup>2,\*</sup>

## Abstract

Electrocardiogram (ECG) signal gives a clear indication whether the heart is at a healthy status or not as the early notification of a cardiac problem in the heart could save the patient's life. Several methods were launched to clarify how to diagnose the abnormality over the ECG signal waves. However, some of them face the problem of lack of accuracy at diagnosis phase of their work. In this research, we present an accurate and successive method for the diagnosis of abnormality through Discrete Wavelet Transform (DWT), QRS complex detection and Support Vector Machines (SVM) classification with overall accuracy rate 95.26%. DWT Refers to sampling any kind of discrete wavelet transform, while SVM is known as a model with related learning algorithm, which is based on supervised learning that perform regression analysis and classification over the data sample. We have tested the ECG signals for 10 patients from different file formats collected from PhysioNet database to observe accuracy level for each patient who needs ECG data to be processed. The results will be presented, in terms of accuracy that ranged from 92.1% to 97.6% and diagnosis status that is classified as either normal or abnormal factors.

**Key Words:** ECG; DWT, SVM.

## I. INTRODUCTION

The current mainstream is highly demanded for a fine wellness status, as the wellbeing of a person measured via the heart, the core part of well-being. The heart status is measured by the produced ECG signal inside it. The heart consists from four main chambers which are simply in a cone muscular pump format spotted in the bore of the chest; exactly located between the lungs and straightway below the sternum [1]. Heart physiology is shown in Fig. 1. Practically, the heart distributes nutrients and oxygen to our tissues; at the same time carrying away carbon dioxide and metabolic waste for excretion through the lungs and the kidneys by pumping the blood in a circulation manner throughout the whole body. The first two chambers located in the upper left and right (called left and right atrium); the third and fourth chambers located in lower left and right are called left and right ventricles. In fact, the senatorial node (SA node), as illustrated in Figure 1 controls the heart contraction rhythm. In addition, it is responsible of triggering the cardiac cycle via the pacemaker by generating electrical pulses and spreading them through the heart conducting through the tissue

pathway.

When the signal moves from SA node through the right of depolarization, through which it moves blood into ventricles [2].

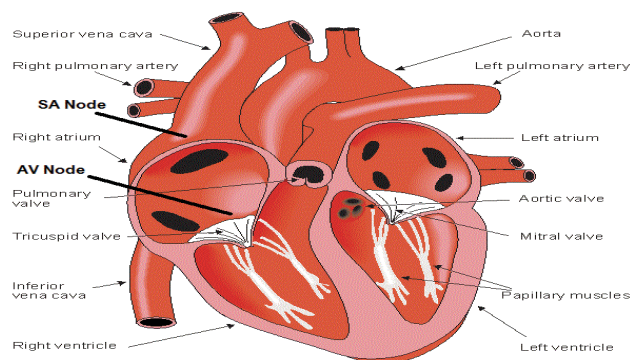


Fig. 1. Heart physiology.

During a normal trial of depolarization, the main electrical vector is directed from SA node towards the AV node, and spreads from the right atrium to the left atrium. This turns into the P wave on the ECG. The PR interval is measured from the beginning of P wave to the beginning of the QRS complex. The PR interval reflects the time of the electrical

Manuscript received June 10, 2016; Revised July 25, 2016; Accepted August 17, 2016. (ID No. JMIS-2016-0010)

Corresponding Author (\*): Hee-Cheol Kim, Department of Computer Engineering, INJE University, Inje-ro 197, Gimhae, Gyeong-Nam, Korea, Tel: 010-8957-6296, E-mail : [heeki@inje.ac.kr](mailto:heeki@inje.ac.kr)

<sup>1</sup>Department of Computer Engineering, INJE University, Gimhae, Korea, E-mail: [ahmadyounes2003@hotmail.com](mailto:ahmadyounes2003@hotmail.com), [joomi@inje.ac.kr](mailto:joomi@inje.ac.kr), [heeki@inje.ac.kr](mailto:heeki@inje.ac.kr)

impulse which travels from the sinus node through the AV node and enters the ventricles; it is normally within 120-200ms. Therefore; the PR interval is a good estimation of AV node function. [3, 4].

In our research, we have proposed a method that diagnoses heart abnormality through Discrete Wavelet Transform (DWT), QRS complex detection and support vector machines (SVM) classification. We have reached a high level of diagnosis accuracy comparing with I. A. Karaye et al, J. A. Delgado et al and T. Tinnakornsriruphap et al methods. The rest of this paper is organized as ECG physiology, related works, contribution and problem stamen, the method, the results and conclusions.

## II. ECG PHYSIOLOGY

Briefly, when the heart cells are in rest situation they considered as in polarized case which means that there is no electrical activity is taking place at the moment. Basically, the heart is depolarized when it is in a systole state and repolarized when it is at diastole state. The depolarization case in normal ECG wave takes place from the beginning the P wave until the middle of ST segment as shown in Figure 2, while the repolarization case starts from the T wave [5, 6].

Generally, ECG features categorized as either waveform or characteristic based features. The easiest one to obtain among the features is characteristic based, because starting extraction process just requests ECG fiducial points knowledge about one of ECG complex. Figure 2 PS, P, PE, Q, R, S, TS, T; T represents the correspondence points that intersect with the peaks and boundaries of P, QRS and T waves, named fiducial points. In ideal ECG complex, there are six boundaries and three peaks (P, R, and T) as illustrated in Figure 2 [7].

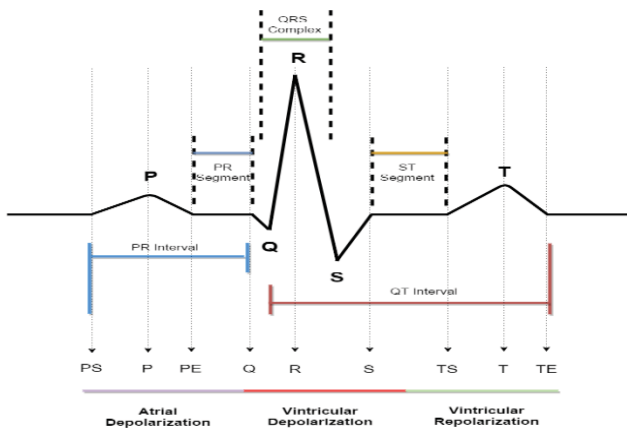


Fig. 2. ECG signal physiology

The main features that aimed to be extracted are angle,

slope, amplitude and duration (temporal). Discriminative characteristics must be selected from these features by applying extra processing over them [8]. The problem of extracting characteristic based features appears when we use fiducial points that may cause difficulties in detecting the boundaries especially in the presence of noise.

Table 1. Normal Duration of waves, intervals and segments in ECG signal.

Feature	Description	Duration
RR Interval	The interval between an R wave and the next R wave: Normal resting heart rate is between 60 and 100 bpm.	0.6 to 1.2s
P Wave	During normal atrial depolarization, the main electrical vector is directed from SA node towards the AV node, and spreads from the right atrium to the left atrium. This turns into the P wave on the ECG.	80m s
PR Interval	The PR interval is measured from the start of P wave to the start of the QRS complex.	120 to 200ms
PR Segment	The PR segment connects P wave and QRS complex. The impulse vector is from the AV node to bundle branches and then to the Purkinje fibers.	50 to 120ms
QRS Complex	The QRS complex reflects the rapid depolarization of the right and the left ventricles. They have a large muscle mass compared to atria, so the QRS complex usually has much larger amplitude than the P-wave.	80 to 120ms
ST Segment	The ST segment connects the QRS complex and T wave. The ST segment represents the period when ventricles are depolarized. It is isoelectric.	80 to 120ms
T Wave	The T wave represents the repolarization (or recovery) of ventricles. The interval from the beginning of QRS complex to the apex of the T wave is referred to as the absolute refractory period.	160ms
ST Interval	It represents the interval between ventricular depolarization and repolarization.	320ms
QT Interval	The QT interval is measured from the beginning of the QRS complex to end of the T wave.	Up to 420ms in heart rate of 60 bpm

On the other hand, Waveform based features processed by using the waveform process coefficient values such as

autocorrelation, fourier coefficient, phase space reconstruction, and wavelet coefficient. It is worth mentioning that more than ECG complexes are used in waveform technique for feature extraction. The technique has an advantage of not requiring detection of the fiducial points which may generate boundary detection difficulties in the presence of noise [9]. The following table shows the normal duration of waves, intervals and segments in the ECG signal.

### III. RELATED WORKS

Some researchers have studied the analysis of ECG signal normal heart rate and abnormal (including right bundle branch block, paced beat, left bundle block branch and atrial premature) beats by using higher order spectral (HOS) method as non-linear tool for complex analysis. They have presented the normal and abnormal signal general characteristics in bispectrum and bicoherence plot to illustrate the difference of visual observation between the normal and abnormal types of beats produced in the heart. Extraction of R-R intervals has been done by using algorithm (pan-Tompkins) and statistical parameters (skewness, kurtosis and variance). The extracted features along with statistical parameters transferred into artificial neural network classifier (ANN). In this work they observed the beats for only 1 minute duration of recording which is not enough; with overall accuracy 94.9% [14].

On the other hand, there is a study through which the researchers they proposed a new method to detect fetal QRS complex from ECG signal of abdominal maternal. They segmented the signal for 250 minutes of contiguous frames; after that they labeled it into 4 classes. They used Principal component analysis over Haar-Wavelet transform to extract the features and decrease the dimensionality. They dealt with an imbalanced class and outliers' problems by applying interquartile ranges and sampling without changing the techniques. By using 10-fold stratified cross-validation of ECG signals 4 electrodes of the PhysioNet/CinC dataset they trained and tested K-nearest neighborhood (KNN), supported vector machine (SVM) and Bayesian network (BN) over the dataset. The resulted accuracy was 89.59% for using KNN and 89.19% for using SVM. But even though KNN resulted as more accurate from SVM, time consuming remains more than SVM. It is worth mentioning that the reached level of accuracy and the duration of time that they tested are not sufficient to obtain the best diagnosis over ECG signals [15].

Another study had been conducted on the basis of medical device communication standard known as Health

Level Seven (HL7). ECG waveform data was represented by ontology and encoding based on HL7. Researchers diagnosed 37 different cardiac abnormalities by an automated diagnosis applying HL7 standard ECG data descriptions and cardiac diagnosis rules. They analyzed ECG waveform data in different formats and used different devices and inside different computer platforms. Therefore, the ontology of ECG waveform data encoded into XML format, thence the collected ECG data from different devices would be acceptable to be displayed in common Internet browser. Their experimental phase intended to examine the system ability of sharing the ECG data in a web browser without errors and measure the diagnosis model accuracy level. They achieved zero error in sharing process with total accuracy of 93% for abnormality diagnosis [16].

### IV. CONTRIBUTION and PROBLEM STATEMENT

Our main contribution is represented in accuracy rate for the diagnosis process of ECG signal with longer duration of time that has significantly increased by applying our method in comparison with the previous related works' methods.

Some previous studies indicated reasonable accuracy results but their methods were applied over short signal duration of time. On the contrary, there were some studies indicated a low accuracy resulted from processing convenient signal duration of time.

### V. METHOD

There are several methods of diagnosing the ECG signals whether it is normal or not, we will present QRS complex detection along with SVM classifier method which is shown as successful and accurate in ECG abnormality diagnosis.

#### 5.1 ECG Records

We have used ten ECG signals records taken from MITBIH arrhythmia database that contained different heartbeat types required for the analysis. The MIT-BIH database that we used contains 48 ECG records. Each record has duration of 30 minutes with sampling frequency of 360 Hz. These records were selected from 24 hours recordings for 47 different individuals. The ECG file formats were with the extensions (\*.mat, \*.dat and \*.txt).

#### 5.2 Pre-processing

Electromagnetic field causes the problem of ECG signal power-line interference, it produces an amount of noise between 50 to 60 Hz. Preprocessing phase includes the elimination of DC component, high frequency noise (by using low pass filter), low frequency noise (by using high pass filter), power line interference by using comb filter, Derivative operation, Squaring operation, Integrator and Thresholding [18].

### 5.3 QRS detection

Detecting the QRS complex is considered as substantial part of the ECG signal analysis. We have used R-peak as an indicator point to extract the QRS complex. In fact, QRS complex is a critical feature among the rest of features in extraction process. Detecting QRS complex was done by using Wavelet energy histogram algorithm. From the phase of QRS detection RR, interval features were obtained for each beat. The determination of complexes within QRS feature and separation of normal and abnormal beats rely on templates matching based on wavelet transformation and QRS detection method [10].

Technically, we used the Matlab function “findpeaks.m” to allocate successfully each peak location, height, width and area. Some peaks are strongly overlapped so we used the Matlab function “peakfit.m”. After locating the peaks positions and heights and eliminating the artifacts, we used the Discrete Wavelet Transform (DWT) for feature extraction step along with QRS threshold for the diagnosis step.

### 5.4 QRS Threshold

We designed our thresholding algorithm by following B. S. Shaik *et.al* [19]; their proposed algorithm has two phases which are training and detection phases. The researchers applied their algorithm successfully over the acquired ECG data from MIT-BIH arrhythmia database. To find out the missed QRS complex during the training we set the recent average and latest R-R interval to 8 beats, signal threshold value to 0.05 and the noise threshold value to  $10^{-3}$ . The thresholding algorithm shows 99.59 % of sensitivity and 99.51% of specificity.

### 5.5 Method diagram

The used method in this study starts by receiving the raw ECG signal which will trigger the whole diagnosis operation. Immediately after receiving the raw ECG data; pre-processing step will take place. After eliminating the noise and artifact the feature extraction process will begin using discrete wavelet transform (DWT) and then divided the data into two data sets training and testing data sets as shown in Figure 3. Before we do the classification phase, the classifier must be trained to imitate the classification

process rapidly. Lastly, the SVM classifier will classify the processed data to normal and abnormal ones.

For non-stationary, transients and aperiodic signals DWT are very powerful and useful to analyze QRS by using Wavelet Energy Histogram Method. In fact, the nature of all biomedical signals (ECG, EEG, ... etc.) are non-stationary hence DWT is required to observe such type of signals. Moreover, analyzing the ECG signal by using DWT, wavelet type and number of decomposition levels selection are conclusive. QRS complex has exotic shape and due to this the energy of signal is different during the QRS complex occurrence. Because of the transition from the Q point to R point and back to S point, energy change occurred. Decomposing the signal with DWT at a suitable level allows picking up the change in energy [10, 11]. A window (segment) of the signal has to be considered for decomposition matter. Finding absolute value is needed because thresholding becomes difficult to be applied since differentiated waveform has bipolar value.

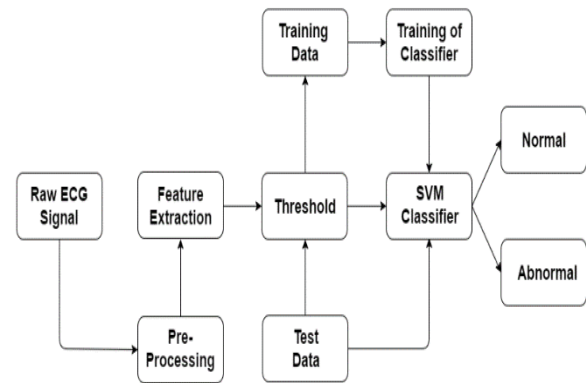


Fig.3. Block diagram.

### 5.6 SVM

In classification phase, we used binary decision tree for a nonlinear SVM based on Radial Basis Function (RBF) to avoid multiclass Classification snag. The selection of kernel function and its parameter values are definitive by using SVM classifier. It is important to optimize the sigma ( $\sigma$ ) value along with other parameters for the kernel function. Using this architecture; (N - 1) SVMs are needed to be trained for an N class problem, but only (log<sub>2</sub> N) SVMs are most required to classify a sample [12, 13]. The proposed classifier architecture SVM is which solves a 2-class classification whether they are normal or abnormal class.

## VI. RESULT

There are several methods of diagnosing the ECG



signals whether it is normal or not, we will present QRS complex detection along with SVM classifier method which is shown as successful and accurate in ECG abnormality diagnosis.

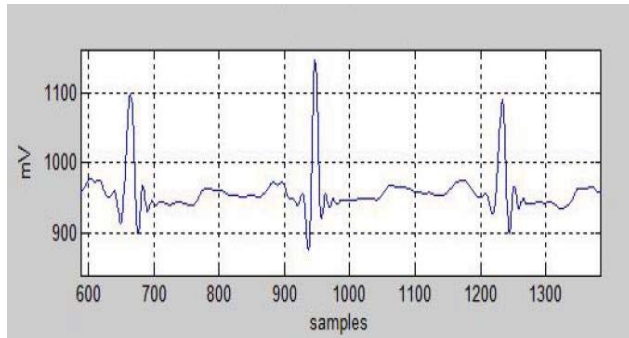


Fig.4. Normal ECG signal.

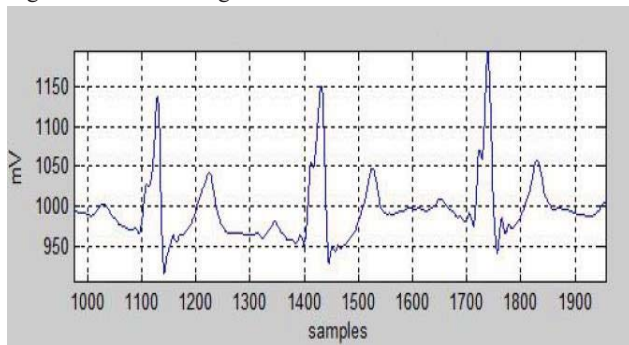


Fig.5. Abnormal ECG signal.

Table.2. ECG records diagnosis results

Record number	Accuracy (Percentage %)	status
01	96.4	normal
02	89.9	normal
03	95.1	normal
04	97.3	abnormal
05	92.1	normal
06	95.5	abnormal
07	97.2	abnormal
08	97.6	normal
09	96.8	abnormal
10	94.7	normal

For calculating the accuracy level of each patient ECG signal, we used the Matlab function “peakfitVSfindpeaks.m”. The overall accuracy rate is 95.26%.

The software that we developed in Matlab, reads the raw Ecg signal, pre-processes it, extracts the features, does thresholding, detects the ectopic beats [17] classifies them into normal and abnormal using SVM. The

following figures present the detected peaks of normal ECG signal from the record 01 as shown in figure 6, figure 7 shows spectrum of normal ECG from the same record and the normal heart rate detection information from the signal and spectrum attached with ectopic beats detection result as shown in figure 8.

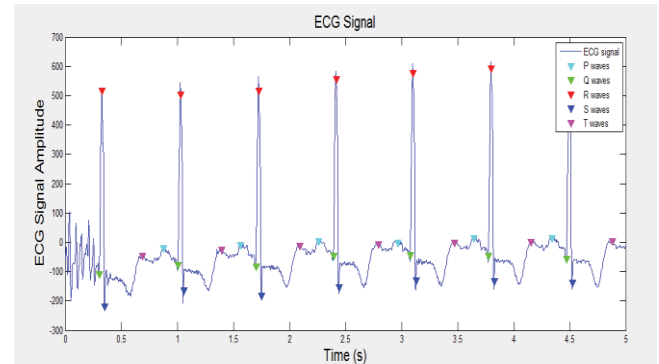


Fig.6. Normal ECG signal with Peaks detection.

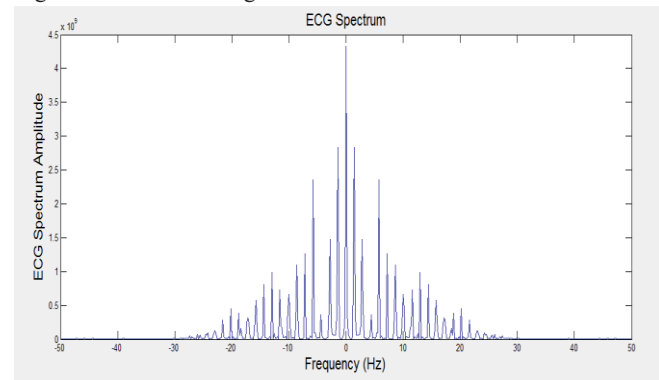


Fig.7. Normal ECG Spectrum.

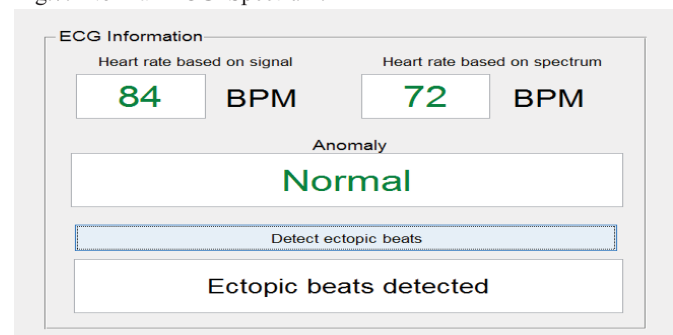


Fig.8. Normal ECG Signal Heart Rate Based on the Signal and Spectrum.

The following figures illustrates the abnormal signal, their peaks detection, spectrum of abnormal ECG and abnormal heart rate detection information along with ectopic beats detection results as shown in figures 9, 10 and 11, respectively.

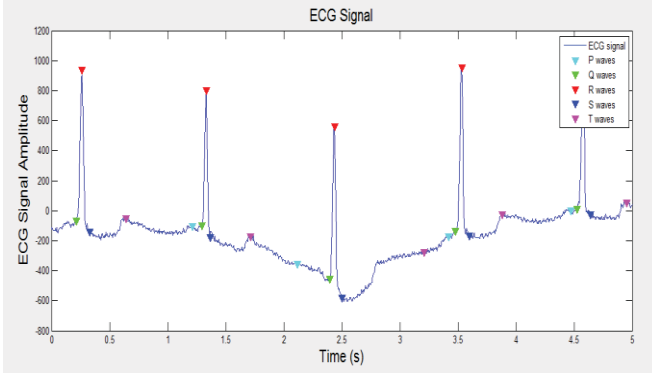


Fig.9. Abnormal ECG signal with Peaks detection.

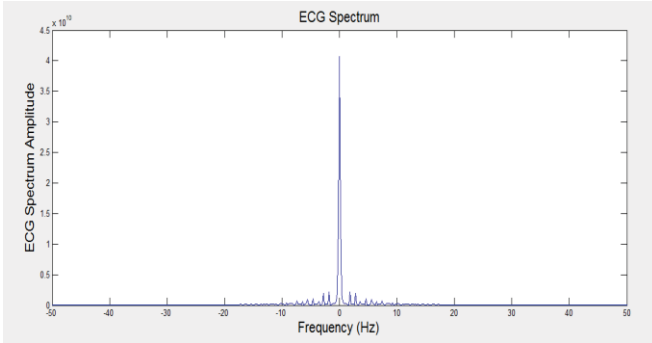


Fig.10. Abnormal ECG Spectrum.

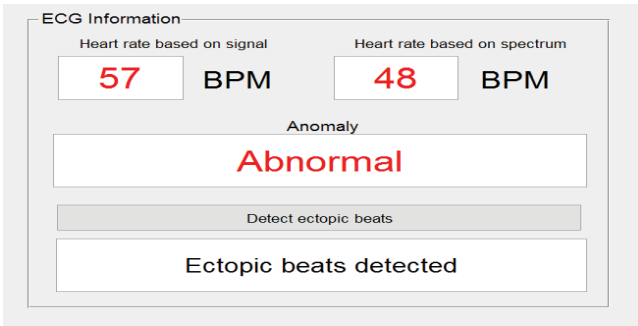


Fig.11. Abnormal ECG signal heart rate based on the signal and spectrum.

Table.3. diagnosis accuracy level comparison between different methods.

Method	Overall accuracy
I. A. Karaye et al.	94.90%
J. A. Delgado et al.	89.59%, 89.19%
T. Tinnakornsriruphap et al	93%
Proposed method	95.26%

## VI. CONCLUSION

The ECG signals of MITBIH arrhythmia database have been analyzed and classified into two main classes: normal or abnormal by using SVM classifier. Depending on the results we get from testing patients' ECG records, we attain a solid derivation that indicates developing diagnosis system, using SVM based on wavelet energy histogram technique associated with DWT for QRS

detection, and which is able to obtain a preferred high and accurate classification rate. The extracted features are consistent and classifiable. Also we have observed that classifier performance depends on the sigma value optimal choice.

## REFERENCES

- [1] G. Xiong, P. Sun, H. Zhou, S. Ha, B. Hartaigh, Q. Truong, and J. Min, "Comprehensive Modeling and Visualization of Cardiac Anatomy and Physiology from CT Imaging and Computer Simulations," in *IEEE Transactions on Visualization and Computer Graphics*, DOI: 10.1109/TVCG.2016.2520946, 2016.
- [2] M. Carvalho, R. Garcês, C. Tavares, I. Rocha, and T. Russomano, "Wavelet analysis of heart rate variability during microgravity simulation," *Bioengineering (ENBENG)*, 2015 *IEEE 4th Portuguese Meeting on*, Porto, 2015.
- [3] M. Bernat and Z. Piotrowski, "Software tool for the analysis of components characteristic for ECG signal," *Mixed Design of Integrated Circuits & Systems (MIXDES)*, 2015 *22nd International Conference*, Torun, pp. 104-109, 2015.
- [4] R. Kumar and A. Kumar, "ECG signal compression algorithm based on joint-multiresolution analysis (J-MRA)," *Electronics and Communication Systems (ICECS)*, 2015 *2nd International Conference on*, Coimbatore, pp. 618-624, 2015.
- [5] K. S. S. Sujana, R. S. Pridhvi, K. P. Priya and R. V. Ramana, "Performance analysis for the Feature Extraction algorithm of an ECG signal," *Innovations in Information, Embedded and Communication Systems (ICIIECS)*, 2015 *International Conference on*, Coimbatore, pp. 1-5, 2015.
- [6] D. Awasthi and S. Madhe, "Analysis of encrypted ECG signal in steganography using wavelet transforms," *Electronics and Communication Systems (ICECS)*, 2015 *2nd International Conference on*, Coimbatore, pp. 718-723, 2015.
- [7] P. Valluraiah and B. Biswal, "ECG signal analysis using Hilbert transform," 2015 *IEEE Power, Communication and Information Technology Conference (PCITC)*, Bhubaneswar, India, pp. 465-469, 2015.
- [8] R. Martinek *et al.*, "Enhanced processing and analysis of multi-channel non-invasive abdominal foetal ECG signals during labor and delivery," in *Electronics Letters*, vol. 51, no. 22, pp. 1744-1746, 2015.
- [9] E. Benmalek and J. Elmhamdi, "Arrhythmia ECG

- signal analysis using non parametric time-frequency technique," Electrical and Information Technologies (ICEIT), 2015 International Conference on, Marrakech, pp. 281-285, 2015.
- [10] W. Jenkal, R. Latif, A. Toumanari, A. Dliou and O. El B'charri, "Enhanced algorithm for QRS detection using discrete wavelet transform (DWT)," *2015 27th International Conference on Microelectronics (ICM)*, Casablanca, pp. 39-42, 2015.
- [11] C. A. Bustamante, S. I. Duque, A. Orozco-Duque and J. Bustamante, "ECG delineation and ischemic ST-segment detection based in wavelet transform and support vector machines," *Health Care Exchanges (PAHCE), 2013 Pan American*, pp. 1-7, 2013.
- [12] S. Faziludeen and P. V. Sabiq, "ECG beat classification using wavelets and SVM," *Information & Communication Technologies (ICT), 2013 IEEE Conference on*, JeJu Island, pp. 815-818, 2013.
- [13] U. Desai, R. J. Martis, C. G. Nayak, Sarika K. and G. Seshikala, "Machine intelligent diagnosis of ECG for arrhythmia classification using DWT, ICA and SVM techniques," *2015 Annual IEEE India Conference (INDICON)*, New Delhi, pp. 1-4, 2015.
- [14] I. A. Karaye, S. Saminu and N. Özkurt, "Analysis of cardiac beats using higher order spectra," *2014 IEEE 6th International Conference on Adaptive Science & Technology (ICAST)*, Ottawa, pp. 1-8, 2014.
- [15] J. A. Delgado, M. Altuve and M. N. Homsí, "Haar wavelet transform and principal component analysis for fetal QRS classification from abdominal maternal ECG recordings," *Signal Processing, Images and Computer Vision (STSIVA), 2015 20th Symposium on*, Bogota, pp. 1-6, 2015.
- [16] T. Tinnakornsrisuphap and R. E. Billo, "An Interoperable System for Automated Diagnosis of Cardiac Abnormalities from Electrocardiogram Data," in *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 2, pp. 493-500, March 2015.
- [17] A. Kennedy, D. D. Finlay, D. Guldenring, R. Bond and J. McLaughlin, "The accuracy of beat-interval based algorithms for detecting atrial fibrillation," *2015 Computing in Cardiology Conference (CinC)*, Nice, pp. 893-896, 2015.
- [18] S. Saminu, N. Özkurt and I. A. Karaye, "Wavelet feature extraction for ECG beat classification," *2014 IEEE 6th International Conference on Adaptive Science & Technology (ICAST)*, Ottawa, pp. 1-6, 2014.
- [19] B. S. Shaik, G. V. S. S. K. R. Naganjaneyulu, and A. V. Narasimhadhan, "A novel approach for QRS delineation in ECG signal based on chirplet transform," *Electronics, Computing and*

*Communication Technologies (CONECCT), 2015 IEEE International Conference on*, Bangalore, pp. 1-5, 2015.

## Authors



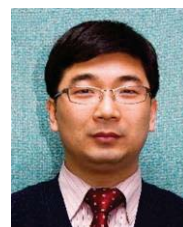
**Ahmed Shdefat** received the B.E. degree in computer information system from Yarmouk University, Jordan, in 2008 and the M.E. degree in information technology from university Utara Malaysia (UUM), Malaysia, in 2010. His previous position was a university lecturer at Computer

Department- College of community service –Taibah University, KSA and currently studies for Ph.D. at Department of Computer Engineering, Inje University, Korea. He has interests in the areas of bio signals, software engineering, and data mining. He has published 3 papers in bio signals.



**Moon-Il Joo** received the MS degree in computer engineering from Inje University in 2012. He is currently a doctor's course at the department of computer engineering, Inje University, Korea. His research interests are in software engineering, human computer interaction, platform based big data, and

component based development.



**Hee-Cheol Kim** received the MS degree in computer science from SoGang University, Korea, in 1991, and the PhD in computer science from Stockholm University, Sweden in 2001. He is a professor at Department of Computer Engineering, Inje University, Korea.

He has interests in the areas of human computer interaction, software engineering, and u-healthcare. He has published more than 70 papers in these areas.

