Research on improving accuracy of Cardiac Disorder data analysis based on Random Forest classifier

HyunJu Lee[1], DongIl Shin[1] and Dongkyoo Shin[1]
Department of Computer Engineering
Sejong University
1. e-mail: nedkelly@gce.sejong.ac.kr, {dshin, shindk}@sejong.ac.kr

HeeWon Park[2] and SooHan Kim[2]
Visual Display Div. R&D Team
SAMSUNG Electronics Co. HQ
2. e-mail:{heewonpark, ksoohan}@samsung.com

***Abstract*— In order to prove that the improved RF algorithm had higher accuracy, the comparing analysis was conducted adapting ECG data. In pre-processing stage, Band-pass Filter was adapted among Wavelet transform, Median Filter, Finite impulse response and others. As a result, the modified Random Forest classifier showed increased more accuracy than SVM, MLP and other researchers' results. Thus, continuous studies on the selection of the filters and methods, which can efficiently delete baseline-wandering at pre-processing phase and accurately extract R-R interval, should be taken place.**

Keywords-ECG; R-R interval; HRV; SVM; MLP; Random Forest; classifier; accuracy.

I. INTRODUCTION

ECG (Electrocardiogram) is an electric signals released by heart activities, which is used as a reference that can identify conditions and diseases of the heart [1]. ECG consists of five ripple marks; P, Q, R, S and T, which verify signals according to height of ripple marks and features of interval, and also can compose ECG data through decision making whether disease exist or not. There is arrhythmia which can be detected by ECG signals, which generally means irregularly fast and slow blood beats [2]. There is MIT-BIH Arrhythmia Database which published for research on arrhythmia.

Signals of ECG are generally experimented based on R-R interval and QRS-Complex extracted data from ECG. Tsipouras, Fotiadis, and Siderise [3] detected and classified arrhythmia according to generated features of heart beat from R-R interval signals. Firstly, they detected signals with blood beats verifying from arrhythmia signals, and then, arrhythmia extraction tasks were secondly conducted with six features released from arrhythmia signals. SVM and MLP classifiers are the most frequently used on ECG experiments. Asl [4], who experimented HRV, proceeded the experiment by two ways; GDA (Generalized Discriminant Analysis) method which is Dimension reducing method was applied into one case of the experiment and GDA was not adapted in another case.

However, it is necessary that experiments on the performance of Random Forest classifier which has differing algorithm compared to SVM and MLP are needed to improve accuracy on experiment results in arrhythmia. Thus, in this study, comparative analysis on accuracies between SVM and MLP classifier was conducted to find out performance of Random Forest classifier. In addition, comparative analysis between parallel data of other researches which experimented with R-R interval extracting and results of this study was also undertaken. R-R interval signal data were verified and constructed, drawing on beat annotation provided by MIT-BIH Arrhythmia Database, and also, modifications of classifier algorithm were attempted.

In this study, there are three different contents in each paragraph state below:

The explanation related to data as well as the process of the experiments was represented in the Section 2. Then, the explanation of the algorithm and the results were commonly noticed in the Section 3. Finally, the conclusion of this study and the direction of further researches were recorded in the Section 4.

II. RELATED WORKS

Meanwhile, there were a lot of experiment concerned with ECG signals and have been applied various filters and classification algorithms. In the case of filters, there were Chazal's [5], Michael's [6], Martinez's [7] works, and so on. Chazal experimented with median filter [5], Michael tested with FIR (finite impulse response) [6], and Martinez tested with wavelet transform [7], however, we experimented with the band-pass filter like Markovsky [8], Taouli [9], and Gholam-Hosseini [10], the band pass filter was judged to be superior to the others and efficient to distinguish the wavelets of ECG by separating whether narrow or wide wavelet.

In the case of classification algorithms, most of which were generally SVM (Support Vector Machine), MLP (Multilayer Perceptron), and DT (Decision Tree), Chau [11], Asl [4], and Bsoul et. al. [12] experimented with SVM and Zhang [13] with the combination of PCA (Principal Characteristics Analysis) and SVM. Also, Inan [14], Yaghouby [15], and Ozbay [16] tested with MLP and Quinlan [17] and Exarchos [18] with DT. And also, Mahesh [19] experimented with Random Forest, Logistic Model Tree, and MLP in classifying the cardiac diseases. Now-a-day, it has come up to more than 90% of accuracy in classifying ECG signals, This paper try to research another method to obtain more accurate rate of classification than existing ones by using the revised Random Forest classifier.

III. DATA AND PRE-PROCESSING

A. MIT-BIH Arrhythmia Database

MIT-BIH (The Massachusetts Institute of Technology – Beth Israel Hospital) Arrhythmia Database [20] is a researched data related arrhythmia analysis with supports receiving from Boston's Israel Hospital and MIT since 1975. MIT-BIH Arrhythmia Database is the first arrhythmia data which can be universally used to detect and evaluate arrhythmia, and total data records are digitalized records from 360 samples per hour per channel. It is ECG records which had been researched in BIH arrhythmia laboratory between 1975 and 1979, measuring patients' movements such as walking through two channels during 24 hours. The database consist of 48 data: 23 numbers of records which were randomly collected from recorded 4000data sets were selected from 40% of outside patients and 60% of hospitalized patients. And other 25 numbers of data included significant arrhythmia signals in clinic although the data were collected from the same patients group.

B. Feature extraction of R-R interval

R-R interval means time of R wave in a human's brain from one certain peak to a next peak, and each R-R interval consists of one cardiac cycle. Fig. 1 indicates R-R interval [21].

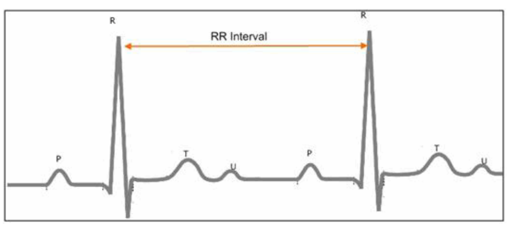

Figure 1. A sample image of R-R interval

R-R interval is continuously generated as a form of continuous time, which is repeated. Sequence of R-R interval are transformed when QRS detector is applied in to ECG signals [22].

Sequences of R-R interval are constituted through time succession, and each sequence which corresponds to immediate heart proportion is defined by the below formula [22].

$$F_i = 1 / RR_i \qquad (1)$$

In general, HRV analyzes HRV in extracted R-R interval using HRV Analysis and constructs HRV data based on analysis information of the extracted HRV. HRV is distinguished into below properties Mean, RMSSD, SDNN, SDSD, NN50, pNN10, pNN5 and so on. In this study, the data properties were classified into total 25 categories including Mean, RMSSD, SDNN, pNN50 and others.

- Mean: inquiring meaning of the 32 number of R-R interval values in each segment.
- RMSSD: meaning the average value of RMS (Root Mean Square) among gaps of intervals from R-R interval.
- SDNN: meaning standard deviation of the gap of R-R interval.
- pNN50: meaning proportions from total section in cases that the gap of R-R interval is over 50cm.

Fig. 2 indicates the feature extraction of R-R interval, and Fig. 3 illustrate HRV analysis. The filter is not only used to delete unnecessary components (frequency components), but also exchange measured data; distances, speeds, accelerations, temperature and strengths, into electric signals. For example, there are Median Filter, finite impulse response, Wavelet transform, Fourier transform and Band-pass Filter, which function as the device (stated above).

In this study's experiments, since Biomedical Startup Kit 3.0 provided by NI LABVIEW (National Instrument LABVIEW) was applied in extraction tasks, Band-pass Filter provided by the kit was adopted. (Fig. 4) Band-pass Filter was designed to filter noises with combining low-pass and high-pass in a single filter [23].

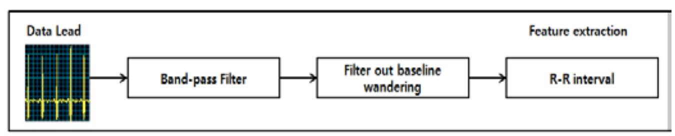

Figure 2. Feature extracion in R-R interval

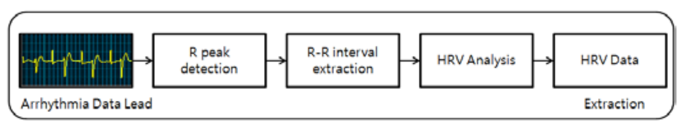

Figure 3. HRV analysis

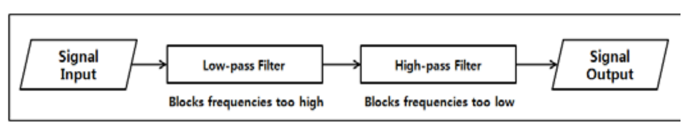

Figure 4. Input and output modes of Band-pass Filter

High-pass and low-pass of the filter were configured at 25Hz and 10Hz respectively. Configured filters erased noises of the data signals, and extracted R-R interval through deleted signals. Then, R-R interval was designed a form to be experimented by WEKA which was used in the classifier experiment. Finally, designed data were experimented by Random Forest [25] classifier, which was one of the classifiers provided by WEKA. Fig. 5 and 6 indicate arrhythmia data before feature extraction and after the extraction. Extracted signals were classified into normal signals and arrhythmia signals according to their intervals and heights. RF is an algorithm belonged to 'tree'. The accuracy of RF was reinforced compared to AF (Atiral Fibrillation) in [27] which had been compared in this study.

And the experiment was performed in [4] with SVM and MLP algorithms applied. In terms of the accuracy of the result, RF relatively showed a higher performance. Therefore, the experiments were undertaken based on RF and the algorithm was also modified to improve the accuracy in this study.

Figure 5. Arrhythmia data before feature extraction

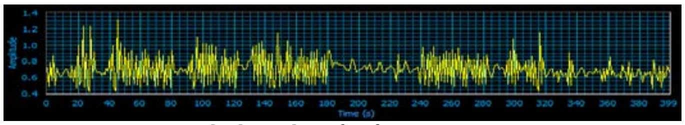

Figure 6. Arrhythmia data after feature extraction

IV. BODY

A. *Modified algorithm of Random Forest classifier and formula*

1) *Modified algorithm of Random Forest*

In 1998, through a research, Ho noticed that Random Subspace is a method to select randomly from each tree with grown subsets using. And after a year, Breiman used new analysis data which was designed to extract results randomly in the original analysis data [24]. Random Forest which is an algorithm that selects random vectors is a specially designed ensemble technique for Decision Tree classifier [25]. Each Decision Tree uses random vectors created from certain possibility distributions. When the tree grow, Decision Tree defines random vectors to segregate each node from selected input features of F numbers rather than totally investigates input features of F numbers selected randomly [25]. It has a input feature called Forest-RI and Forest-RC: Forest-RI is a way which randomly select a vector of the RI, Forest-RC divides input data into the beat condition when input features of F numbers reach universal linear compounding [25]. In modified algorithm in this study, Forest-RI was designed to select the most frequent signals and Forest-RC was designed to classify arrhythmia chased by the algorithm. And Best-First decision tree (B-F tree) was applied rather than Decision Tree.

In general, Decision Tree which classifies target variables has had an aim that classified a given data. Also, selecting the most related variables and target variables, tree compounds categories and separates the most related category, which tree has a limitation according to features of the basic data.

Thus, tree cannot guarantee the best accuracy because tree becomes too sophisticated and the rate of the classificati on shows low performance when the features of the data are not vertically classified to certain variables. Therefore, to complement these drawbacks, B-F Tree is applied to modify algorithm. B-F Tree which is a method that extends nodes with best-fit order rather than fixed orders minimizes errors which come from all nodes needing separation with the most efficiently separated nodes added in each experiment stage. In each stage, tree extends with the most modified subset selecting. And the constructed process is expanded when all of the nodes reach a certain number or a pure node [26]. At a stage of pruning, the first B-F Tree can conduct two methods; pre-pruning and post-pruning.

When tree is growing, pre-pruning stops its growth, if data, which are divided, are not practical. The second post-pruning, which continuously extends nodes until all of the trees are completely extended, and it selects with extended data numbers and sorts of the average error estimates-minimizing [26]. Two cases were made based on all data of the final tree and extended selecting numbers.

2) *Formula*

Through experimenting Random Forest classifier, Accuracy, Sensitivity, Specificity and PPV (Positive Predictive Value) were measured after TP (True Positive), TN (True Negative), FP (False Positive) and FN (False Negative) had been extracted. The formula is stated below.

$$\text{Accuracy} = TP + TN / TP + TN + FP + FN \quad (2)$$
$$\text{Sensitivity} = TP / TP + FN \quad (3)$$
$$\text{Specificity} = TN / TN + FP \quad (4)$$
$$\text{PPV} = TP / TP + FP \quad (5)$$

If results of the formulas are needed to exchange into percentage, each results of the formulas is just multiplied 100.

B. *The method of the experiment*

In this study's experiments, R-R interval was extracted as two wave forms; Narrow and Wide with Band-pass filter using. In constructed arrhythmia data, N (Normal) and ~ (Change in signal quality) signals were the most frequent beats: N means the regular wave form and ~ reveals that wave forms become to shift their current form. V (Premature Ventricular Construction) and A (Atrial Premature Beat) signals, which are commonly mean arrhythmia, had been generated the most frequently in this study's experiments, therefore, the algorithm was modified based on those the most frequent beats. In the modified algorithm, Forest-RI preferentially chose N (Normal) and ~ (Change in signal quality) signals, and Forest-RC were designed to classify V (Premature Ventricular Constraction) and A (Atrial Premature Beat) signals after chasing them. And then, after N and ~ signals were separated, Forest-RI was constructed to chase and classify other signals as well. The modified Random Forest algorithm is stated below.

- Forest-RI firstly chases F = N (Normal) and F = ~ (Change in signal quality), then verifies them.
    - o N = Normal signal
    - o ~ = altering signals // Forest-RI means input selection
- Forest-RI chases V (Premature Ventricular Constraction) and A (Atrial Premature Beat)
- Forest-RC distributes V and A into arrhythmia // Forest-RC means the highest separation
    - o V = Arrhythmia
    - o A = Arrhythmia
- Forest-RI chases other signals and distributes, excepting N, ~, V and A signals
- Forest-RI and Forest-RC are repeated
- Forest-RI / Forest-RC are ended

In this study, apart from the modification of Forest-RI and Forest-RC, Beat-First decision tree (B-F Tree) was applied rather than Decision Tree in order to reduce out-of-bag. B-F Tree, which uses best-fit order to extend nodes without fixed orders, stops its growth otherwise segregated data do not show actuality, and it makes decisions with finally extended data volumes adapting. Thus, it can decrease out-of-bag rates more easily than Decision Tree as minimizing extended volumes and branches. When Forest-RI and Forest-RC had been able to chase and classify selectively, TP (True Positive) showed relatively high values before its modification, while values of FP (False Positive) were decreased. And out-of-bag was relatively reduced compared to the past experiments when B-F Tree had been applied. Therefore, its accuracy remarkably higher than the results of other established experiments. A selected datum was experimented to identify that the performance of B-F Tree was more excellent than Decision Tree. From the result, two facts stated below were revealed. The accuracy of Decision Tree was 90.69%: its volume and leave were 341 and 171 respectively, whereas the accuracy of B-F Tree was 93.37% as its volume and leave of tree were 567 and 284 respectively. This study's actual classifier experiments were conducted with Random Forest classifier in WEKA–3.6.2 version Fig. 7. In the experiments, R-R interval had been extracted at the pre-processing stage, and extracted data were then corrected to be experimented by WEKA. Experiments using WEKA had been firstly compressed, Random Forest classifier experimented the compressed data. The experiment using Random Forest was undertaken by k-fold-cross-validation method. Seperating data as k number of times of the same size section, k-fold-cross-validation method is a means that, an experimental section is selected among other sections and the others are used as training materials [25]. According to these sequences, each section is repeated in order to be used exactly once only, and total out-of-bags added by k number of times of the total experiments. In this study, configuring k to 10, experiments had been performed by using 10-fold-cross-validation, and then, Accuracy, Sensitivity, Specificity and PPV (Positive Predictable Value) were measured based on extracted figures of TP, TN, FP and FN. And Accuracy, Sensitivity and Specificity were just measured from feature data of HRV in this study.

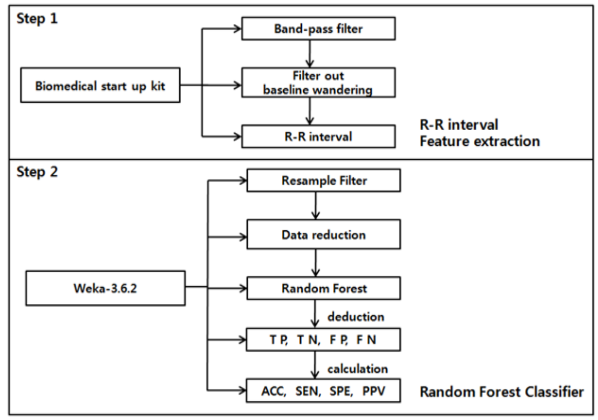

Figure 7. The process of the experiments

C. The results of the experiment

1) Accuracy comparison on the modification of the algorithm of Random Forest classifier between pre-results and now

In this study, an algorithm of RF (Random Forest) was modified to chase preferentially selected signals in order to improve its accuracy of results. Fig. 8 indicates differences between before its modification and after.

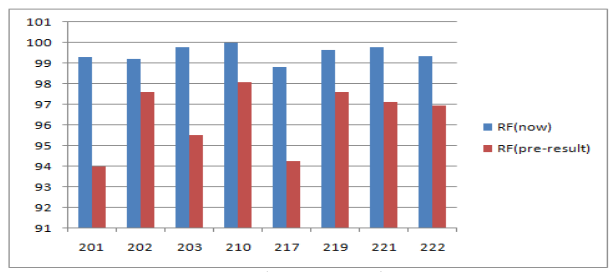

Figure 8. Accuracy comparison of the results of Random Forest algorithm modification between before and after

Data were selected based on consequences of K&L's experiment [27]. The sequence of the results stated at the Fig 8 is from RF (now) to RF (pre-result): RF (now) is results of the experiment after the modification, RF (pre-result) is results before the modification. From the Fig. 8, the algorithm after the modification shows more accurate than its counterpart before the modification.

2) Accuracy comparision among established researches, SVM, MLP and Random Forest

Table 1 indicates extracted results of Sensitivity, Specificity and Accuracy on R-R interval experiments which were conducted by established researches and Random Forest classifier.

TABLE 1. MEASUREMENT OF SENSITIVITY (SEN), SPECIFICITY (SPE) AND ACCURACY (ACC)

Arrhythmia Data		Sensitivity(%)		Specificity(%)		Accuracy(%)	
		K&L	*RF*	*K&L*	*RF*	*K&L*	*RF*
Record_ no	*201*	96.44	99.71	39.52	88.23	65.46	99.28
	202	80.79	99.62	94.64	72.73	88.72	99.2
	203	81.36	99.89	21.92	71.43	63.37	99.77
	210	96.89	100	0	0	94.73	100
	217	72.78	99.51	94.22	27.27	90.86	98.81
	219	96.86	99.81	64.20	89.66	91.39	99.63
	221	92.25	99.86	50.66	94.87	65.24	99.76
	222	92.25	100	50.66	0	65.24	100
Average		90.226	99.75	61	63.1	82.14	99.55

TABLE 3. ACCURACY ANALYSIS AMONG RF, MLP AND SVM

Record_no	Accuracy(%)		
	RF	*MLP*	*SVM*
201	99.28	83.69	81.27
202	99.2	96.42	96.23
203	99.77	80.42	77
210	100	96.38	95.69
217	98.81	75.95	68.45
219	99.63	92.89	92.89
221	99.76	86.36	82.31
222	100	77.34	74.87
Average	99.55	86.18	83.58

TABLE 2. RESULTS ON FEATURE DATA OF R-R INTERVAL EXPERIMENTED BY RANDOM FOREST CLASSIFIER

Record	TP	TN	FP	FN	PPV	SEN	SPE	ACC
201	1748	60	8	5	99.54	99.71	88.23	**99.28**
202	2081	24	9	8	99.56	99.62	72.73	**99.2**
203	2972	10	4	3	99.86	99.89	71.43	**99.77**
210	2352	25	0	0	100	100	0	**100**
217	2246	6	16	11	99.29	99.51	27.27	**98.81**
219	1600	26	3	3	99.81	99.81	89.66	**99.63**
221	2066	37	2	3	99.9	99.86	94.87	**99.76**
222	2567	0	0	0	100	100	0	**100**
Average	2204	23.5	5.25	4.125	99.745	99.8	55.52	**99.55**

Table 1 indicates the results of K&L, RF: K&L is Tateno's experiment [27], RF is this study's experiment. When it comes to comparison among them, it is stated below:

While the result of the 210 sector is the best in K&L's experiment, RF shows better performance as high as 5.27% compared to K&L.

Table 2 represents results that 8 number of feature data were experimented by Random Forest (RF).

After values of TP, TN, FP and FN had been previously extracted, PPV (Positive Prediction Value percentage), SEN (Sensitivity percentage), SPE (Specificity percentage) and ACC (Accuracy percentage) were measured, and then, the Average was calculated. Including those data, all of the other data showed over 90% of accuracy rates as well.

Thus, it could be regarede that Random Forest classifier extracted efficient Accuracy in the results of total data.

In order to analyze the accuracy of RF, Table 3 shows Accuracy rates among SVM, MLP, and RF. The sequence of the table is in order; RF -> MLP -> SVM. Through Table 3, it could be obviously comprehended that the accuracy of RF reaches approximately 100% compared to others.

3) Results cpmparison of HRV experiments between this study and established researches

Asl [4], which is the comparison of HRV (Heart Rate Variability) experiments on HRV, was used SVM (Support Vector Machine) and MLP (Multilayer Perceptron), and conducted by two ways; one was the case that GDA (Generalized Discriminant Analysis) which is 'Dimension reduce' method was adapted into the experiment and another was the 'GDA' was not adapted (no GDA) in common.

Thus, in this study, experiments were undertaken by two ways called 'All data' and 'Shorten' methods: total 25 number of properties were used in 'All data' method and the only 13 number of properties were used in 'Shorten' method. And Random Forest, MLP and SVM were commonly selected as algorithms of the experiments. Results of the experiments were then differently compared by cases: the results of 'All data' was compared with 'no GDA' [4] and its counterpart of 'Shorten' was compared with 'GDA'.

From the results of the experiment, 'GDA' shows better performance on 'Accuracy (ACC)' than 'no GDA' on HRV and SVM at 0.27% and 0.67% respectively. From the case of this study's consequence, 'Shorten' method indicated higher 'Accuracy' on MLP and SVM at 1.05% and 3.12% respectively.

TABLE 4. Results comparison of the experiment on HRV

Method		SEN (%)	SPE (%)	ACC (%)	Average
MLP	No GDA	90.64	98.51	98.22	95.79
	GDA	92.63	98.98	98.49	96.7
	Shorten	100	90.9	98.96	96.62
	All data	100	83.33	97.91	93.746
SVM	No GDA	92.57	98.88	98.49	96.646
	GDA	95.77	99.4	99.16	98.11
	Shorten	100	66.67	97.91	88.2
	All data	100	83.33	94.79	92.7
RF	Shorten	100	90.9	98.96	96.62
	All data	100	90.9	98.96	96.62

In terms of the comparison between 'GDA' and 'Shorten' method, although its 'Accuracy' was high at 0.47% when 'Shorten' method had adapted MLP, its 'Accuracy' was low at 1.25% when SVM was used. Finally, when RF (Random Forest) was adapted into the experiment, their 'Accuracy' (between the cases of MLP and SVM) were at 98.96% at the same time. Why this study's results did not show remarkably higher performance than Asl [4] could be assumed that there was the lack of efficiency in the experiments of this study compared to Asl [4]'s research on 'pre-processing' as well as 'Dimension reduction' of the data.

V. CONCLUSION AND DIRECTION OF CONTINUOUS STUDY

In this study, the Accuracy rates among SVM, MLP and Random Forest classifiers were adapted into the comparative analysis of their performances using MIT-BIH Arrhythmia Database. Biomedical Startup Kit used the data extraction of R-R interval at pre-processing phases and Random Forest classifier provided by WEKA_was used with its algorithm modified. In order to emphasize differences between the two groups, the algorithm of Random Forest classifier was modified by below three steps:

- The algorithm was changed to select high-frequent signals previously instead of random selection
- The algorithm was corrected to detect arrhythmia signals in the best-fitted segregation and classify them.
- Best-First decision tree was applied instead of Decision Tree.

As a result, the accuracy of Random Forest classifier could be remarkably maximized, and classifier did not show only higher performance than SVM and MLP classifier, but could also minimize out-of-bags. And, it was proved that the modified algorithm presented higher accuracy rates compared with the results of K. Tateno's researches in the aspect of the accuracy.

Consequently, despite of that remarkably high results were gained on the improvement of 10% accuracy rate, there were lower results at pre-processing phase than B. M. Asl's research process in terms of the next areas; exceeding limitation on Dimension reduction and used Band-pass Filter as well as efficient section separation of R-R interval. Therefore, after this study, it should be researched to select filter which can efficiently erase baseline-wandering in pre-processing phase and investigate methods that can accurately extract R-R interval.

ACKNOWLEDGMENT

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2009.

REFERENCES

[1] 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 Classification of ECG Beat Using Higher Order Statistics and Hermite Model," *K Kor Soc Med Informatics*, vol. 15, pp. 117-131, 2009.

[2] Korean Geart Rhythm Society. [Online]. Available (April 13, 2012): http://www.k-hrs.org/

[3] M. G. Tsipouras, D. I. Fotiadis , and D. Sideris, "An arrhythmia Classification system based on the RR-interval signal," *Artificial Intelligence in Madicine*, vol. 33, pp. 237-250, 2005.

[4] B. M. Asl, S. K. Setarehdan, and M. Mohebbi, "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal," *Artificial Intelligence in Medicine*, vol. 44, pp. 51-64, 2008.

[5] P. D. Chazal, M. O' Dwyer, and R. B. Reilly, "Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features," *IEEE Transactions on Biomedical Engineering*, vol. 51, no.7, pp. 1196-1206, 2004.

[6] Michael and L. Hilton, "Wavelet and Wavelet Packet Compression of Electrocardiograms," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 5, pp. 394-402, 1997.

[7] J. P. Martinez, R. Almeida, and S. Olmos, "A Wavelet-Based ECG Delineator: Evaluation on Standard Databases," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 4, pp. 570-581, 2004.

[8] I. Markovsky, A. Amann, and S. V. Huffel, "Application of Filtering Methods for Removal of Resuscitation Artifacts from Human ECG Signals," *30th Annual International IEEE EMBS Conference*, pp. 13-16, 2008.

[9] S. A. Taouli and F. B. Reguig, "Detection of QRS Complexes in ECG Signals Based on Empirical Mode Decomposition," *Global Journal of Computer Science and Technology*, vol. 11, pp. 11-17, 2011.

[10] H. Gholam-Hosseini, H. Nazeran, and K. J. Reynoids, "ECG Noise Cancellation Using Digital Filters," *2nd International Conference on Bioelectromagnetism*, pp. 151-152, 1998.

[11] K.C. Chua, V. Chandran, R.U. Acharya, L.C. Min, "Cardiac Health Diagnosis Using Higher Order Spectra and Support Vector Machine," *Open Med Inform J 3*, pp. 1–8, 2009.

[12] M. Bsoul et al., "Real-time sleep quality assessment using single-lead ECG and multi-stage SVM classifier," *in the International Conference of IEEE Engineering in Medicine and Biology Society*, pp. 1178–1181, 2010.

[13] H. Zhang and L. Q. Zhang, "ECG analysis based on PCA and Support Vector Machines," *Networks and Brain, IEEE*, pp. 743-747, 2005.

[14] O. T. Inan and G. T. A. Kovacs, "Robust Neural-Network-Based Classification of Premature Ventricular Contractions Using Wavelet Transform and Timing Interval Features," *IEEE transctions on biomedical engineering*, vol. 53, no. 12, 2006.

[15] F. Yaghouby, A. Ayatollahi, R. Soleimani, "Classification of Cardiac Abnormalities Using Reduced Features of Heart Rate Variability Signal," *World Applied Sciences Journal*, vol. 6, no. 11, pp. 1547–1554, 2009.

[16] Y. Özbay, R. Ceylan and B. Karlik, "A fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Comput. Biol. Med.*, vol. 36, pp. 376–388, 2006.

[17] J.R. Quinlan, "C4.5: Programs for machine learning," *Morgan Kaufmann Publishers, San Francisco CA*, 1993.

[18] T.P. Exarchos, M.G. Tsipouras, C.P. Exarchos, C. Papaloukas, D.I. Fotiadis and L.K. Michalis, "A methodology for the automated creation of fuzzy expert systems for ischaemic and arrhythmic beat classification based on a set of rules obtained by a decision tree," *Artif. Intell. Med.*, vol. 40, no. 3, pp. 187–200, 2007.

[19] V. Mahesh, A. Kandaswamy, C. Vimal and B. Sathish, "Cardiac disease classification using heart rate signals," *Int. J. Electronic Healthcare*, vol. 5, no. 3, 2010.

[20] PhysioBank. [Online]. Available (April 13, 2012): http://physionet.mit.edu/physiobank/database/mitdb/

[21] NI Biomedical Startup Kit 3.0. [Online]. Available (April 13, 2012): http://decibel.ni.com/content/docs/DOC-12646

[22] G. D. Clifford, F. Azuaje, and P. E. McSharry, “Advanced Methods and Tools for ECG Data Analysis,” *Artech House*, pp. 101-102, 2006.

[23] All About Circuits URL. [Online]. Aailable (April 13, 2012): http://www.allaboutcircuits.com/

[24] L. Breiman, “Machine Learing,” *Kluwer Academic Publishers in Netherlands*, 2001.

[25] P. N. Tan, M. Steinbach and V. Kumar, “Introduction to Data Mining,” *1-st Addison-Weseley,* 2006.

[26] H. Shi, “Best-first Decision Tree Learning,” *The University of Waikato in NewZealand,* pp. 3-5, 2007.

[27] K. Tateno and L. Glass, “A Method for Decision of Atrial Fibillation Using RR-Intervlas,” *Computers in Cardiology(IEEE),* vol. 27, pp. 391-394, 2000.