# On the use of Auto-Regressive Modeling for Arrhythmia Detection

Mourad Adnane
Electronics Engineering Department/LDCCP Lab
Ecole Nationale Polytechnique
Algiers, Algeria 16200
Email: mourad.adnane@enp.edu.dz

Adel Belouchrani
Electrical Engineering Department/LDCCP Lab
Ecole Nationale Polytechnique
Algiers, Algeria 16200
Email: adel.belouchrani@enp.edu.dz

***Abstract*—This paper investigates the use of an auto-regressive modeling method for the classification of heartbeats into two categories: Normal beats (N) and Ventricular ectopic beats (VEB). The method is based on an auto-regressive modeling (AR) of QRS complexes. Each heartbeat is characterized by its AR coefficients. Then, K-nearest neighbor (K-NN) classifier uses the AR coefficients to discriminate between N beats and VEB. In addition, the use of AR modeling prediction error $e_n$ as a discriminating feature is investigated. Results show that the prediction error power ($\sigma_p^2$) enhances significantly the classification accuracy. The proposed classifier is compared to a classifier based on the use of RR timing information. Finally, the two classifiers are combined together where the classification result is given by the agreement of the two classifiers. The proposed AR modeling approach performs better than the RR interval-based classifier and their combination enhances the classification accuracy.**

## I. Introduction

Conventional way for the classification of heartbeats in medical environment, which is based on visual inspection, is tedious and time consuming. Researchers are always proposing different methods of automatic classification in order to solve these problems and to offer to medical practiotionners computer aided diagnosis tools. However, the characteristics of heartbeats varying from an individual to another and from a time to another time for the same individual makes this task difficult. Indeed, there is no common criterion or feature to be used for the classification of all kinds of heartbeats. This is why the research in this field is very wide. Many algorithms were proposed in the literature and the main difference between them concerns the feature selection/extraction method. Herein are mentioned some notable research papers and the methods applied to extract features: In [1] authors used 26 features based on ECG morphology and RR interval information. In [2] same authors improve their method by making their algorithm adaptable by using a fraction of the local record used in the training step. This method, however, is hardly applicable since a fraction of each patient's ECG should be acquired and used in the training step. A combination of several time and time-frequency features are used in [3]. In [4] several time features were extracted and used in mixture of expert system classifier. It is to be noticed that a user-specific training data set is used in this paper. Research in [5] shows that wavelet coefficients can be used to discriminate between normal beats and arrhythmias. This principle is applied in [6] to detect VEB. Authors in [7] used only one method to extract features which is the auto-regressive modeling of ECG cycles. The authors used a fourth order AR model and used the AR coefficients in a linear classifier. The AR coefficients were computed using Burg's algorithm. Research in the field of heartbeats classification is very wide and a good review of the proposed methods in the literature can be found in [8]. In this paper, we propose a simple method based on the use of AR modeling method where only QRS complexes are modeled. The AR coefficients are computed using Levinson algorithm. The noise level (prediction error) power estimated by the Levinson algorithm is also used as a feature as it shows that it improves classification rate. A comparison with a classifier based on the use of RR interval information which proves to be one of the best arrhythmia characterizing features is done. Finally, the two classifiers are combined together. The method was tested against 6 hours data obtained from the MIT/BIH Arrhythmia database (MITDB) [9] and very satisfactory results were obtained which proves the validity of the proposed approach. This paper is organized as follows: in section 2, the proposed method is described in detail. In section 3, results are presented and a discussion is made. Finally, a conclusion is given in section 4.

## II. Method

The American National Standards Institute (ANSI) and the Association for the Advancement of Medical Instrumentation (AAMI), released the ANSI/AAMI EC57:1998 standard [10] which recommend the use of five beat classes such as normal beat, ventricular ectopic beat (VEB), supraventricular ectopic beat (SVEB), fusion of a normal and a VEB, or unknown beat type. Usually, the supraventricular ectopic beat (SVEB) category is merged with normal beat category and labeled as normal (N) in a first classification step. This is the case in [4] for instance. This is not in contradiction with AAMI standard. A classifier can separate, in a first step, between two categories and then make other separations inside each category to reach the five beat classes. In this paper, we use two categories: beats labeled normal (N) including SVEB category and ventricular ectopic beats (VEB). Actually, these categories represent the majority of beats. For instance, the

978-0-9928-6265-7/16/$31.00 ©2016 IEEE 

TABLE I
MITDB RECORDS USED IN TRAINING AND TESTING STEPS

| Records | N beats | VEB | Total beats |
|---|---|---|---|
| Training Data | | | |
| 100 | 2271 | 1 | 2272 |
| 105 | 2525 | 41 | 2566 |
| 111 | 2122 | 1 | 2123 |
| 114 | 1831 | 43 | 1874 |
| 116 | 2302 | 109 | 2411 |
| 119 | 1542 | 444 | 1986 |
| 121 | 1861 | 1 | 1862 |
| Total | 14454 | 640 | 15094 |
| Testing Data | | | |
| 208 | 1587 | 992 | 2579 |
| 209 | 3003 | 1 | 3004 |
| 221 | 2030 | 396 | 2426 |
| 230 | 2254 | 1 | 2255 |
| 234 | 2749 | 3 | 2752 |
| Total | 11623 | 1393 | 13016 |
| Total data | 26077 | 2033 | 28110 |

MIT/BIH arrhythmia database (MITDB) [9] which is very popular among ECG processing algorithm developers, include 100825 beats in the aforementioned five beats categories (paced beats are not included in this number). Only 0.03 % (33 beats) are labeled as unclassified and only 0.8 % (803 beats) are labeled as fusion of a normal and a VEB. We can say without loss of generality, that developing an algorithm treating only the three categories (N), (VEB) and (SVEB) can be easily generalized to all kinds of beats since these categories represent the majority of beats (99.17 % of beats in the MITDB). Then, in this work we use two categories: N (including normal and SVEB categories) and VEB category. In the following, N stands for beats labeled as normal (Normal and SVEB) and VEB stands for ventricular ectopic beats.

### A. MIT/BIH arrhythmia database (MITDB)

MIT/BIH arrhythmia database (MITDB) contains 48 ECG records of 30 minutes each. Records contain a variety of normal and arrhythmia ECG beats. MITDB is extensively used in automatic heartbeat (QRS) detection algorithms as well as in automatic heartbeat classification algorithms. Sampling frequency is 360 Hz with 11 bits depth. Every record is accompanied with an annotation file containing beats labels. In this paper, one ECG lead is used. 12 records containing N beats and VEB are used for training and testing steps such as: 7 records for training and 5 records for testing. Detail about the beats used in training and testing steps is given in Table I.

### B. Heartbeat classification Method

The proposed heartbeat classification algorithm is divided into three steps: preprocessing step, feature extraction step and finally, classifier training and testing step. Details of each step are given subsequently.

### C. Preprocessing

The raw ECG signal is corrupted with several artifacts among which are: baseline wander, EMG noise, power line noise, etc. It is important to reduce the effects of these artifacts before any processing of the data. Herein, we used a cascade of three filters: a 3rd-order Butterworth high-pass filter (HPF) with a cut-off frequency of 1 Hz to eliminate baseline wander and DC component. The second filter is a 3rd-order Butterworth band rejection filter (BRF) used to eliminate 60 Hz power line interference and finally, a 4th-order Butterworth low-pass filter (LPF) with a cut-off frequency of 25 Hz to eliminate high frequency artifacts such as EMG noise.

### D. Feature extraction

The purpose behind feature extraction is to characterize data (in our case ECG) by some measures which can be used to discriminate between different states (categories). The choice of the kind of measures to use is a critical point. For heartbeat classification purpose, two ECG characteristics show great potential in the discrimination between normal and arrhythmia beats. The first one is the QRS morphology and the second one is the timing information (distance between successive R waves, R to T waves distance, etc.). Figure 1 shows an example of normal beats (represented with N in Fig. 1) and 1 ventricular ectopic beat (shown by an arrow in Fig. 1). We can see clearly that the magnitudes and widths of VEB and normal beats in Figure 1 are different. Similarly, we can see that VEB's pre-RR and post-RR interval values are different from those of normal beats. Then, it appears judicious to choose morphology features and timing information features in any heartbeat classification scheme. In this paper, we use two kinds of features, one related to the morphology (the proposed AR modeling method) and another one related to timing information (used for comparison purpose). These two kinds of features are explained subsequently.

### E. AR modeling

We use an AR modeling of QRS complexes. Each beat will be represented by some AR coefficients (depending on AR model order). First of all, we use the annotation file of the MITDB [9] giving the R peak positions within each ECG record to determine the QRS complex. It should be noticed that normal QRS duration is generally smaller than 120 ms [11]. This duration corresponds to 43.2 samples in the MITDB. Also, duration of QRS complex corresponding to arrhythmia beats is often greater than 120 ms and can reach 150 ms [11] which correspond to 54 samples. We have chosen to use 60 samples surrounding the position of each R peak to get the QRS complexes in each case (29 samples before R peak position and 30 samples after R peak position). Each QRS complex is modeled using AR model as the output of a linear recursive system with a white noise as an input:

$$x_n = e_n + \sum_{i=1}^{P} -a_i x_{n-i} \tag{1}$$

Where $x_n$ is the QRS complex sequence ($n = 60$ in this paper), $e_n$ is the white noise, $a_i$ are the AR coefficients and $P$ indicates AR model order. The summation term in equation (1) is a linear prediction of $x_n$ based on previous samples. Therefore, $e_n$ represents the error of prediction. The best linear

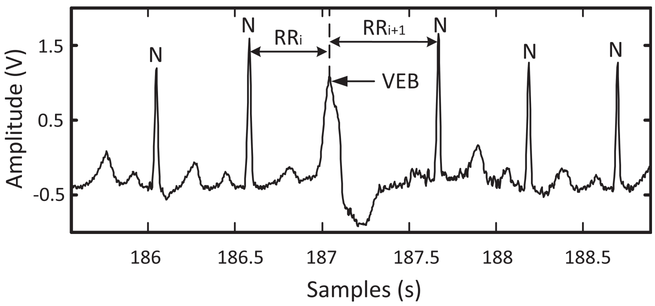


Fig. 1. Example of heartbeats from the record 208. N depicts Normal beats. An arrow depicts VEB. VEB's Pre-RR and Post-RR intervals are shown.

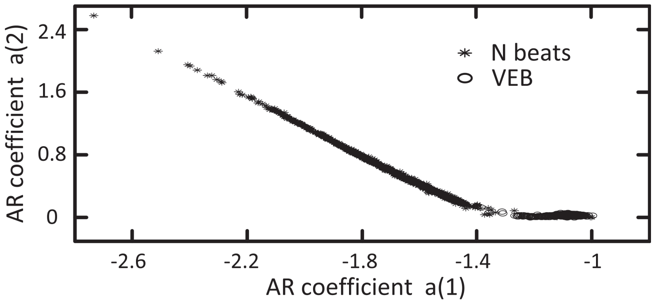


Fig. 2. AR coefficient $a_2$ versus AR coefficient $a_1$ corresponding to beats of the record 208. Asterisks depict N beats whereas circles depict VEB.

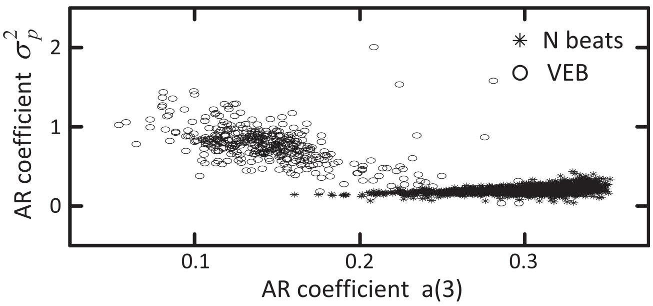


Fig. 3. Prediction error power $\sigma_p^2$ versus AR coefficient $a_3$ corresponding to beats of the record 221. Asterisks depict N beats whereas circles depict VEB.

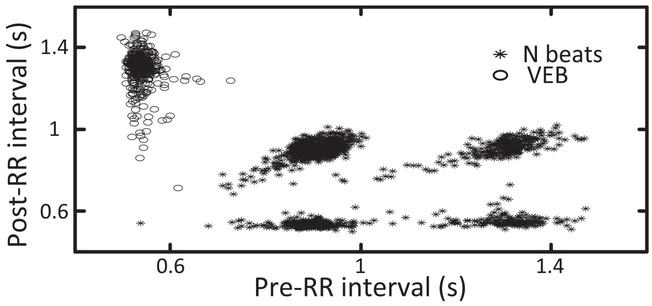


Fig. 4. Post-RR interval versus Pre-RR interval corresponding to beats of the record 119. Asterisks depict N beats whereas circles depict VEB.

prediction is obtained by minimizing error of prediction. The modeling of equation (1) leads to the Yule-Walker equations stated as: (for more details see [12])

$$\underline{r} = R\underline{a} \tag{2}$$

Where $\underline{r}$ is the column vector of auto-correlations, $R$ the matrix of auto-correlations and $\underline{a}$ the column vector of AR coefficients. To determine $\underline{a}$, it is sufficient to inverse the matrix $R$ but this is very costly in term of calculation. Then, a fast algorithm to determine $\underline{a}$ is used. This algorithm is the Levinson recursion algorithm (for details refer to [12]). In this paper, we use the AR coefficients and the noise level power (prediction error power) expressed by $\sigma_p^2$ obtained for a model order $P = 3$, which means there are 4 features obtained from AR modeling associated to each beat, such as ($a_1$, $a_2$, $a_3$, $\sigma_p^2$). Figure 2 shows an example of the coefficients $a_2$ versus $a_1$ corresponding to the ECG beats of the record number 208 of the MITDB. An example of the coefficients $\sigma_p^2$ versus $a_3$ corresponding to the ECG beats of the record number 221 of the MITDB is depicted in Figure 3. Figures 2 and 3 show clearly the potential of AR coefficients to discriminate between N and VEB categories.

### *F. RR information*

The RR information is obtained by measuring the distance (time duration) between successive R waves. In this paper, we use two RR values corresponding to each peak. The two RR values correspond to: Pre-RR interval and Post-RR interval. These two features are expressed in seconds. An example of the coefficients Post-RR interval versus Pre-RR interval corresponding to the ECG beats of the record number 119 of the MITDB is depicted in Figure 4.

### *G. Heartbeat classification*

The classification is made in two main steps. First, the training data (see table I) is used. This data is divided into two clusters containing normal beats and VEB respectively. The two clusters are used separately in a k-means clustering algorithm as a single cluster entry in each case. This operation is made to calculate each cluster centroid location. Actually, k-means is usually used to partition $m$ samples of a given data into $k$ clusters. The main idea behind k-means is to minimize the total sum of distances over the $k$ clusters and within each cluster. The samples belong to the cluster for which the distance towards its centroid location is minimal in comparison to other clusters centroid locations. For more details about k-means method refer to [13].

In this paper, k-means is used exclusively to calculate centroid locations for N beats cluster and VEB cluster. This operation applied to the training data gives the centroid locations for N beats cluster and VEB cluster as shown in Table II. The second step of the classification is made using testing data and a k-NN classifier [14]. k-NN clasifier or k-nearest neighbor algorithm principle is to classify each beat by comparing its entry (features) to a known-class object. Basically, given the N beats cluster centroids locations and VEB cluster centroids

TABLE II
CENTROID LOCATIONS FOR N BEATS AND VEB

| | $a_1$ | $a_2$ | $a_3$ | $\sigma_p^2$ | Pre-RR | Post-RR |
|---|---|---|---|---|---|---|
| N cluster Centroids | -1.6025 | 0.4205 | 0.2367 | 0.1425 | 0.8222 | 0.8083 |
| VEB cluster Centroids | -1.2578 | 0.0221 | 0.2284 | 1.6774 | 0.5319 | 1.2944 |

obtained in first step, any entry (features) of the testing data (features corresponding to a given beat) are compared to these two cluster centroids. If the tested beat features are closer to the N beats cluster centroid location, it will be classified as normal beat. Similarly, if the tested beat features are closer to the VEB cluster centroid location, it will be classified as VEB. The cluster centroids obtained in first step are divided into RR related centroids and AR related centroids. Then, two classifiers are built, one based on the AR modeling features and the AR centroids (columns 2 to 5 of Table II). The second one is based on the RR interval features and the RR interval centroids values (columns 6 and 7 of Table II). Each classifier classifies the training data separately. Finally, another classifier is built based on the combination of the two aforementioned classifiers where a heartbeat is labeled N or VEB only if the two classifiers agreed, i.e., both classifiers labeled the beat as normal (N) or both of them labeled it VEB. Beats which are not classified similarly by the two classifiers are rejected (unclassified). This is done to decrease misclassifications and to get high confidence about the classified heartbeats.

## III. RESULTS AND DISCUSSION

In order to evaluate the performance of our algorithm, we tested it using 12 records of the MITDB. Five performance metrics were used to evaluate our algorithm: the detection accuracy ($Acc$), the sensitivity ($Se$), the specificity ($Sp$), the positive predictivity ($Pp$) and rejected beats ratio ($F$) :

$$Acc = \frac{TP + TN}{TP + FN + TN + FP} \times 100 \quad (3)$$

$$Se = \frac{TP}{TP + FN} \times 100 \quad (4)$$

$$Sp = \frac{TN}{TN + FP} \times 100 \quad (5)$$

$$Pp = \frac{TP}{TP + FP} \times 100 \quad (6)$$

$$F = \frac{Rb}{Tb} \times 100 \quad (7)$$

Where TP (true positives) is the number of VEB correctly classified, FN (false negatives) is the number of VEB wrongly classified (classified as N beats), TN (true negatives) is the number of N beats correctly classified, FP (false positives) is the number of N beats wrongly classified (classified as VEB). Detection accuracy evaluates the overall accuracy of the algorithm. The sensitivity measures the ability of the algorithm to detect VEB. The specificity measures the ability of the algorithm to detect N beats. Positive predictivity measures the fraction of true VEB among all the beats classified as VEB. $F$ measures the fraction of rejected beats (not classified) in the classifier combining scheme where $Rb$ represents the number of rejected beats and $Tb$ is the total number of beats.

TABLE III
COMPARISON BETWEEN AR MODEL-BASED CLASSIFIER RESULTS AND RR INTERVAL-BASED CLASSIFIER RESULTS

| | Errors | $Acc$ (%) | $Se$ (%) | $Sp$ (%) | $Pp$ (%) |
|---|---|---|---|---|---|
| AR model-based classifier | 90 | 99.31 | 93.61 | 99.99 | 99.92 |
| RR interval-based classifier | 692 | 94.68 | 60.23 | 98.81 | 85.88 |

TABLE IV
INFLUENCE OF $\sigma_p^2$ ON THE AR MODEL-BASED CLASSIFIER PERFORMANCE

| | Errors | $Acc$ (%) | $Se$ (%) | $Sp$ (%) | $Pp$ (%) |
|---|---|---|---|---|---|
| Classifier with $\sigma_p^2$ | 90 | 99.31 | 93.61 | 99.99 | 99.92 |
| Classifier without $\sigma_p^2$ | 741 | 94.30 | 100 | 93.63 | 65.28 |

Classification results obtained for the proposed two classifiers applied to the testing data (presented in Table I) are presented in Table III. It appears clearly that the proposed AR model-based method is performing much better than the RR interval-based method.

In order to show the importance of the feature $\sigma_p^2$ in the classification, we tested the AR classifier once including $\sigma_p^2$ and once taking it off. Results are summarized in Table IV. It appears clearly that the performance results are much better when $\sigma_p^2$ is included. For instance, the detection accuracy reached 99.31% when $\sigma_p^2$ was included and only 94.3% when it was taken off. We notice also that the positive predictivity value is bad in the case $\sigma_p^2$ is taken off (Pp=65.28%) which contrasts with the good result obtained when $\sigma_p^2$ is included (Pp=99.92%). This shows clearly the improvements obtained when adding the $\sigma_p^2$ feature. This result makes sense. Actually, $\sigma_p^2$ represents the prediction error power value. This value is logically bigger in the case of VEB. Recall Figure 1 where examples of normal beat and VEB are shown. It appears when screening this figure that for the same AR model order (say $P = 3$) the normal QRS complex would be better fitted than it would be for a VEB. Then, the prediction error power $\sigma_p^2$ is logically bigger in the case of VEB. However, more investigations should be made to check the impact of noisy ECG signals on the relevancy of $\sigma_p^2$ on a hand and investigations should be made to check the impact of the AR model order on the relevancy of $\sigma_p^2$. We suspect $\sigma_p^2$ to be less relevant in the classification task when the order of the AR model increases or when the noise level increases.

The combining classification scheme is applied to testing data. Good results are observed in term of accuracy, however a quite big rejection ratio is obtained: F=10.67%, 1250 VEB and 139 N beats. A simple method to remedy to that problem is to re-inject the category of unclassified beats into the classifier again, but this time with new clusters centroid values. These cluster values are calculated from the category which was already classified. This means that a portion of a training data is used to classify the other portion of the data. This scheme can be though as a subject-specific scheme but it is not. Actually, in subject-specific scheme a portion of a record

TABLE V
CLASSIFICATION RESULTS OBTAINED FOR THE TESTING DATA IN THE CLASSIFIER COMBINING SCHEME

| Records | Total beats | $Rb$ | $F$ (%) | N beats | V beats | TP | FP | FN | TN | $Acc$ (%) | $Se$ (%) | $Sp$ (%) | $Pp$ (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 208 | 2579 | 222 | 8.6 | 1579 | 778 | 773 | 0 | 5 | 1579 | 99.79 | 99.36 | 100 | 100 |
| 209 | 3004 | 48 | 1.6 | 2956 | 0 | – | 0 | – | 2956 | 100 | – | 100 | – |
| 221 | 2426 | 161 | 6.64 | 1948 | 317 | 313 | 0 | 4 | 1948 | 99.82 | 98.74 | 100 | 100 |
| 230 | 2255 | 1 | 0.04 | 2253 | 1 | 1 | 0 | 0 | 2253 | 100 | 100 | 100 | 100 |
| 234 | 2752 | 3 | 0.11 | 2748 | 1 | 1 | 0 | 0 | 2748 | 100 | 100 | 100 | 100 |
| Total | 13016 | 435 | 3.34 | 11484 | 1097 | 1088 | 0 | 9 | 11484 | 99.93 | 99.18 | 100 | 100 |

with known class is used to classify the other portion of the same record (say same individual). But in our case, the algorithm does not have a prior idea about the real class of the training data. The algorithm classifies a portion of the data and then it uses these results to classify the rest of the data. Overall results are presented in Table V. Results show that the aforementioned method enables to reduce the number of rejected beats significantly. The rejection ratio decreased to 3.34% of total tested heartbeats (139 N beats and 296 VEB). This means that 954 VEB which were rejeted in the first round of classification were classified in the second round of classification and more important, they were correctly classified as it is shown in Table V. Indeed, very good results are reported. The overall detection accuracy reaches 99.93% which shows that combining the two kind of features into one classifier is beneficial. The algorithm demonstrated its ability to detect VEB since sensitivity reaches 99.18%. The ability of the algorithm to detect N beats was also demonstrated since specificity reaches 100%. Another important result relies to the positive predictivity which reaches 100%. This is very important result. Actually, Pp=100% means that all the beats which were classified as VEB are true VEB. Actually, it is always better to reject a heartbeat than to wrongly classify it. Suppose for instance that a portion of the rejected VEB (296 VEB in this paper) is wrongly classified as N beats , say 50% of them. The result will be a big number of VEB classified as normal beats. Therefore, these beats will not be inspected by medical practitioners. Indeed, re-verifying all the beats classified as N beats is very fastidious because the majority of beats are classified as normal and the majority of heartbeats are normal. In contrary to that, it is much easier to check a small portion of the data manually (the 3.34% rejected beats) to check whether they belong to N beats class or VEB class. Many conventional arrhythmia detectors use the morphology of the ECG to discriminate between normal and arrhythmia beats. AR modeling of QRS waves has shown, by its simplicity and efficiency, its ability to be a serious alternative to the conventional morphology features.

## IV. CONCLUSION

In this paper, a simple method of heartbeat classification based on AR modeling of QRS complexes is investigated. The method shows very good results. Among the notable points, the use of the prediction error power $\sigma_p^2$ as a feature enhances significantly the classification accuracy. In addition, we notice that $\sigma_p^2$ value in the case of N beats is much smaller than it is in the case of VEB for the same AR model order.

The proposed method is compared to a classifier based on an extensively used feature which is the RR interval information. Results show that the proposed approach performs better than the RR interval-based approach. The combination of the two classifiers enhances the classification accuracy and allows the algorithm to detect very accurately N beats and VEB and to reject suspicious cases which enhances the confidence in the obtained results. Finally, the proposed method opens the door to research with regards to model choice applied to QRS wave.

## REFERENCES


[1] P. de Chazal, M. O'Dwyer, and R. B. Reily, Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features, *IEEE Trans. Biomed. Eng.*, vol.51, pp.1196-2006, Jul. 2004.

[2] P. de Chazal, and R. B. Reily, A Patient-Adapting Heartbeat Classifier Using ECG Morphology and Heartbeat Interval Features, *IEEE Trans. Biomed. Eng.*, vol.53, pp.2535-2545, Dec. 2006.

[3] M. G. Tsipouras, and D. I. Fotiadis, Automatic arrhythmia detection based on time and timefrequency analysis of heart rate variability, *Comput. Method. Program. Biomed.*, vol.74, pp.95-108, Feb. 2004.

[4] Y. H. Hu, S. Palreddy, and W. J. Tompkins, A Patient-Adaptable ECG Beat Classifier Using a Mixture of Experts Approach, *IEEE Trans. Biomed. Eng.*, vol.44, pp.891-900, Sep. 1997.

[5] Z. Zidelmal, A. Amirou, M. Adnane, and A. Belouchrani, QRS detection based on wavelet coefficients, *Comput. Method. Program. Biomed.*, vol.107, pp.490-496, Mar. 2012.

[6] M. Adnane, and A. Belouchrani, Premature Ventricular Contraction Arrhythmia Detection Using Wavelet Coefficients, *8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA), IEEE Xplore*, pp.170-173, May 2013.

[7] D. Ge, N. Srinivasan, and S. M. Krishnan, Cardiac arrhythmia classification using autoregressive modeling, *Biomed. ENG. Online.*, vol. 1, no.5, Nov. 2002.

[8] R. J. Martis, U. R. Acharya, and H. Adeli, Current methods in electrocardiogram characterization, *Comput. Biol. Med.*, vol.48, pp.133-149, Jan. 2014.

[9] R. Mark, and G. Moody, MIT-BIH arrhythmia database, *http://www.physionet.org/physiobank/database/mitdb/*, accessed: 25 December 2014.

[10] Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms, 1998. Association for the Advancement of Medical lnstrumentation.,

[11] Radford et al., ACC/AHA Heart Failure Clinical Data Standards ACC/AHA Key data elements and definitions for measuring the clinical management and outcomes of patients with chronic heart failure, *J. Am. Coll. Cardiol.* vol.46, no.6, pp.1180-1207, 2005.

[12] S. V. Vaseghi, "Linear Prediction Models," in *Advanced Digital Signal Processing and Noise Reduction*, 2nd ed. New York, Wiley, 2002, ch.8, pp.227-262.

[13] J. MacQueen, Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1967, 1:281-297.

[14] T. M. Cover, and P. E. Hart, Nearest neighbor pattern classification, *IEEE Trans. Inform. Theory.*, vol.13, pp. 21-27, Jan. 1967.