

Detection and Classification of Baseline-Wander Noise in ECG Signals Using Discrete Wavelet Transform and Decision Tree Classifier

Syed Irtaza Haider¹, Musaed Alhussein²

¹*College of Computer and Information Sciences, King Saud University,
Riyadh, Saudi Arabia*

²*Department of Computer Engineering, College of Computer and Information Sciences, King Saud University,*

*Riyadh 11543, Saudi Arabia
sirtaza@ksu.edu.sa*

Abstract—An electrocardiogram (ECG) signal is usually contaminated with various noises, such as baseline-wander, power-line interference, and electromyogram (EMG) noise. Denoising must be performed to extract meaningful information from ECG signals for clinical detection of heart diseases. This work is focused on baseline-wander noise as it shares the same frequency spectrum as the ST segment of ECG signals. Hence, it is important to estimate the baseline-wander prior to its removal from ECG signals. This paper presents a method for classifying each segment of the ECG signal's baseline-wander as minimal, moderate or large. We use the C4.5 decision tree algorithm to model the classifier using the WEKA data-mining tool. We test the proposed method on ECG signals obtained from the MIT-BIH arrhythmia database (48 ECG recordings, each slightly longer than 30 min). We use 36 ECG recordings for training the classifier with the remaining 12 ECG recordings as the test data for classification. We partition each recording into 5 second, non-overlapping segments, which result in 361 segments for each record. The classification results show that the model classifier achieves an average sensitivity of 97.36 %, specificity of 99.50 %, and overall accuracy of 98.89 % in classifying the baseline-wander noise in ECG signals. The proposed method effectively addresses the question of identifying the minimal baseline-wander segments. Moreover, the proposed framework may help in devising an algorithm for the selective filtering of moderate and large baseline-wander segments to achieve the best trade-off between accuracy and computational cost.

Index Terms—Biomedical signal processing; Discrete wavelet transforms; Decision trees; Electrocardiogram; MIT-BIH arrhythmia database; Signal Denoising.

I. INTRODUCTION

Electrocardiogram (ECG) is an important clinical tool that provides cardiac information for the early diagnosis of cardiac diseases. A typical ECG signal consists of a P wave, QRS-complex wave, and T wave [1]. By extracting the morphological features of ECG signals, heart rate variability and cardiac diseases can be detected.

Manuscript received 29 November, 2018; accepted 30 April, 2019.

This work was supported by the Research Center of College of Computer and Information Sciences, King Saud University. The authors are grateful for this support.

ECG signal processing requires research in several areas including data acquisition [2], [3], signal amplification [4], [5], denoising techniques, and morphological and statistical feature extraction. At the data acquisition stage, a low-noise amplifier with high input impedance and high common mode rejection ratio is required to maximize the amplitude of the signal without distorting it.

When the ECG signal is acquired, it is usually contaminated owing to the presence of several noise sources, and thus a pre-processing stage is necessary. Some common noises are baseline-wander, patient-electrode motion artifact, electrode-contact noise, power-line interference, and EMG noise [6]. The pre-processing stage must ensure that the morphological features of the acquired ECG signal are not compromised during denoising in order to improve the signal-to-noise ratio, thus making the signal analysis much more accurate and effective.

Baseline-wander is one of the major low-frequency noises in the range well below 0.8 Hz [7]. It is mainly due to patient's movement, respiration, muscle contraction, and electrode-impedance change [8]. The presence of baseline-wander might cause ECG annotators to inaccurately identify ECG features [9]. Care must be taken while eliminating the baseline-wander as the low-frequency region of the ECG spectrum (ST segment) is highly susceptible to distortion.

In the past, denoising of the ECG baseline-wander has been explored. A number of filtering techniques have been proposed and developed to remove the low-frequency noise from the ECG signal. Among them, high pass finite impulse response (FIR) filters [10] and infinite impulse response (IIR) filters [9] are the most commonly used classical techniques. The American Health Association (AHA) recommends that the filter cut-off frequency should not be more than 0.05 Hz in order to prevent possible distortion of the ST segment, but this frequency can be relaxed to 0.67 Hz as long as the linear phase of the filter is preserved [11], [12]. The problem with FIR filters is that they have higher filter order and memory requirements [13]. In addition, the use of FIR filters may cause longer delays. In contrast, IIR filters have a non-linear phase response that can be resolved

by using a zero phase IIR filter that requires reverse filtering of the entire data. Some other filtering techniques include median filters [9], [14], adaptive filters [7], [15], and moving average filters [8].

Although adaptive filters have good noise elimination characteristics, they distort the low-frequency component of an ECG signal. In order to overcome this limitation, a wavelet adaptive filter is proposed in [16]. Recently, various techniques have been proposed based on wavelet transforms [17]–[20], interpolation algorithms [21], sparse signal decomposition [22], principal component analysis [23], independent component analysis [24], and empirical mode decomposition (EMD) [25] method. EMD is computationally very demanding in comparison to the other techniques [7].

The remainder of this paper is organized as follows. Section III presents the proposed method. The results and discussions are presented in Section IV and Section V, respectively. Finally, the conclusions drawn are presented in Section VI.

II. MOTIVATION

Each baseline-wander removal technique has advantages and drawbacks based on the complexity of the technique and accuracy of the filtered signal. It is known that the ST segment morphology provides useful information for the clinical detection of heart diseases and cardiac stress testing. Hence, it is extremely important to estimate the baseline-wander prior to its removal from the ECG signal. A good estimation of the baseline-wander will prevent the filtering of baseline-free segments of the ECG signal, thus ensuring the accuracy of the acquired signal.

In order to illustrate this idea, an example of an ECG recording corrupted by baseline-wander is shown in Fig. 1(a). This segment is taken from MIT-BIH arrhythmia database record no. 115. From the figure, it can be observed that the signal has large baseline-wander noise, which must be eliminated prior to the signal analysis. Fig. 1(b) shows the ECG recording with a clear ST depression but no baseline-wander. This segment is taken from the European ST-T database record e0106 (Signal: V3). As shown in the figure, it is a baseline-free segment and does not require any filter.

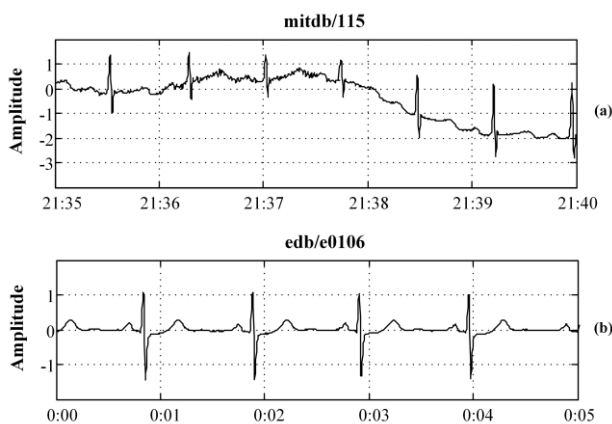


Fig. 1. (a) ECG recording corrupted by baseline-wander noise. This segment of ECG recording is taken from the MIT-BIH arrhythmia database record no. 115 (Signal: MLI2); (b) ECG recording with a clear ST depression, but no baseline-wander. This segment of the ECG recording is taken from the European ST-T database record e0106 (Signal: V3).

The high-pass filter can distort the ST segment of an ECG signal. This is demonstrated using the example in Fig. 2. The original signal is taken from the European ST-T database and contains 6 second of data extracted from record e0106. We applied a second order high-pass Butterworth filter with a cut-off frequency of 0.5 Hz to this baseline-free segment. As shown in the figure, there is a clear ST depression in the original signal, but after applying the filter, the ST depression is reduced in the filtered signal.

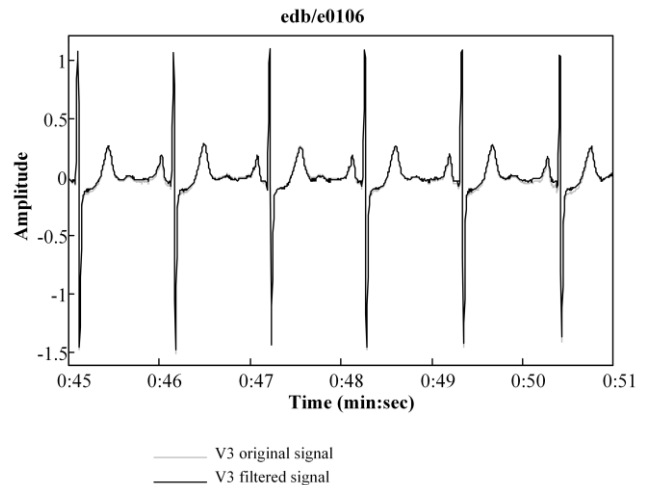


Fig. 2. ECG segment (grey) is taken from the European ST-T database record e0106. Filtered signal (black) obtained on applying second order high-pass Butterworth filter.

In summary, traditional baseline-removal techniques introduce distortion in the ECG signal to some extent even if the signal is baseline-free. Therefore, it is desirable to distinguish between a baseline-free segment and a baseline-wander segment prior to applying any filtering techniques in order to eliminate the baseline-wander from the signal. Our primary motivation for developing the method presented in this paper is to identify the segments of an ECG signal that have minimal or no baseline-wander. This identification process assists in preserving the clinical relevance of minimal or baseline-free segments by not applying any filtering technique. The proposed method can successfully identify minimal baseline-wander segments of ECG signal. Moreover, the proposed method can be employed for the selective filtering of moderate and large baseline-wander segments of ECG signal. This approach is comparatively simple, effective, and computationally undemanding.

III. MATERIALS AND METHODS

The characteristics of an ECG signal become distorted on applying a filtering technique for eliminating baseline-wander noise. Therefore, identification of minimal or baseline-free segments is necessary. Fig. 3 illustrates the framework of the proposed methodology. The main idea is to classify each segment of an ECG signal into minimal, moderate, and large baseline-wander noise. We use the wavelet transform in our study. The proposed algorithm comprises five stages: (a) partition of ECG signal into 5 second non-overlapping segments, (b) down-sampling, (c) wavelet decomposition and reconstruction of approximation coefficients at level 6 using *coif5* as the mother wavelet, (d) magnitude estimation of the baseline-wander noise, and (e)

construction of a decision tree classifier.

A. Segmentation without Overlap

Segmentation is the process of dividing a signal into discrete, non-overlapping fragments. The objective of segmentation is to break the input signal into multiple fragments and analyze each fragment separately. It is often much easier to process a signal by breaking it into segments. It is ideally suited for real-time applications because of its simple implementation and lack of pre-processing.

The segmentation without overlap method is widely used for ECG signal quality classification. A study in [26] discusses the effect of segment size on the performance of the ECG signal quality classification. The results in [26]

show that the performance of the ECG signal quality classification using a 5 and 7 second segment size is good as compared with other segment sizes. In this work, we consider a segment size of 5 seconds.

As shown in Fig. 3, the first step is to divide ECG signal $S(n)$ of each record that is N points long into m equal non-overlapped segments $S_1(L), S_2(L), \dots, S_m(L)$. The length L of each segment S_i , where $i = 1, 2, \dots, m$, is expressed as $L = f_s \times t$, where f_s is the sampling frequency of the ECG signal, and t is the segment size in seconds. The total number of segments m is expressed as $m = \frac{N}{L}$, where N is the length of the ECG signal $S(n)$, and L is the total number of samples in each segment.

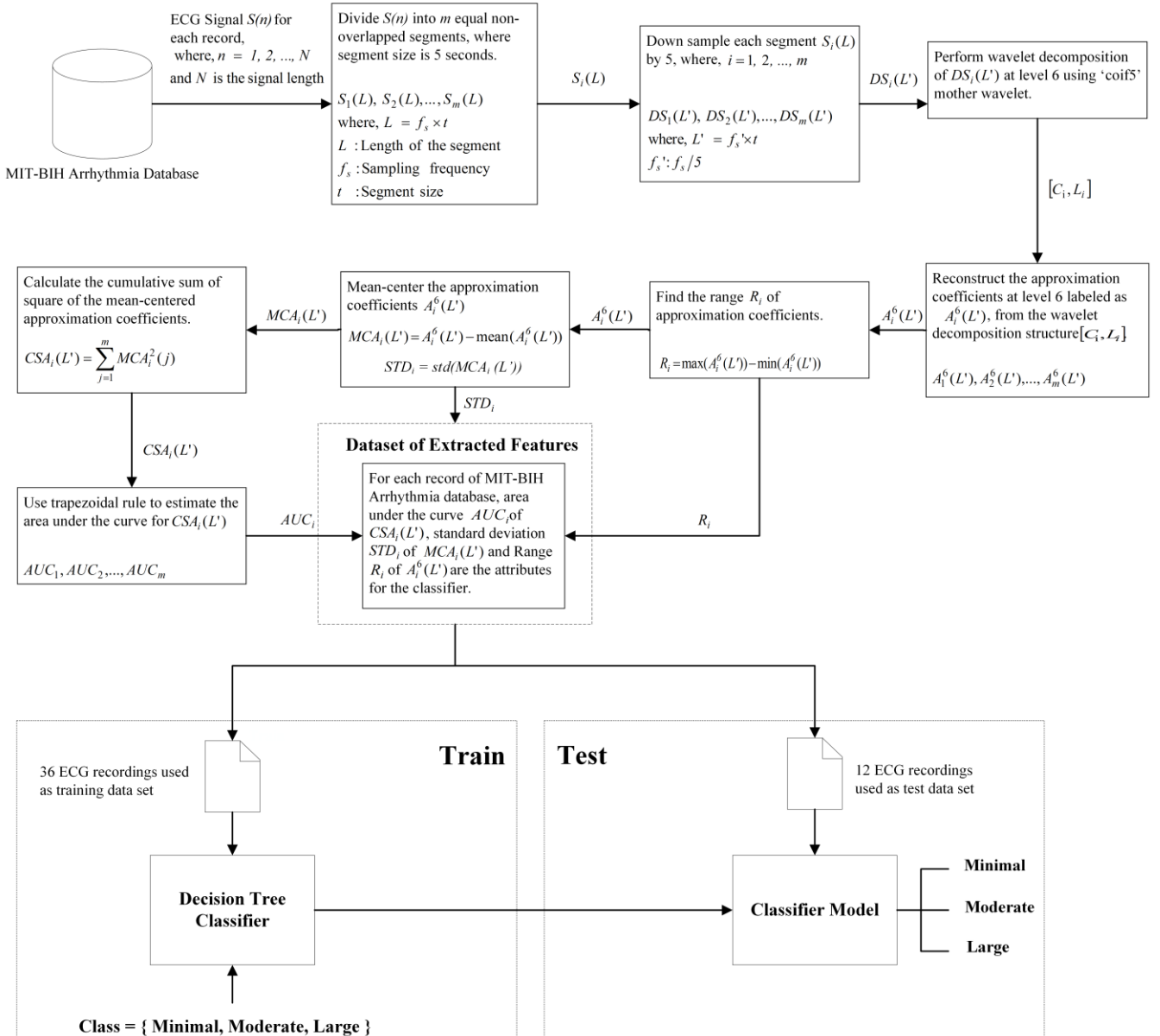


Fig. 3. Proposed methodology to classify baseline-wander noise.

B. Down-Sampling

The next step is the down-sampling of each segment by a factor of 5. Usually an anti-aliasing low-pass filter is applied prior to performing down-sampling, but if the desired frequency is within the range of the down-sampled frequency, the use of the anti-aliasing filter can be avoided.

As most of the ECGs spectral power is located below 30 Hz [27], [28], using the Nyquist sampling theorem, we can down-sample the data to 60 Hz or greater, which is the desired frequency in this case. If the signal is acquired at a sampling frequency of 360 Hz, we can down-sample by a factor of 5 Hz to 72 Hz. This new down-sampled frequency is greater than the desired frequency. Moreover, we are not

interested in the high-frequency contents of the ECG signal as we are estimating the baseline-wander noise, and thus we do not require the anti-aliasing filter.

The down-sampling not only reduces the data bandwidth as well as the computational load, but also reduces the numerical load of the signal processing algorithms.

C. Discrete Wavelet Transformation

The discrete wavelet transform (DWT) is a spectral estimation tool for the time-frequency analysis of non-stationary signals. It is widely used in ECG denoising. The choice of an optimal mother wavelet function and the levels of signal decomposition plays a vital role in estimating the baseline-wander noise.

We recommend two approaches for selecting the optimal levels of signal decomposition using the wavelet transform. First, we use (1)

$$Level = \text{ceil} \left[\log_2 \left(\frac{f_s}{2} \right) \right], \quad (1)$$

where f_s is the sampling frequency of the signal. The second approach to finding the decomposition level for baseline-wander noise suppression is given below

$$Level = \text{ceil} \left[\left(1 - \frac{\log_2 2t}{\log_2 L'} \right) \times \log_2 L' \right], \quad (2)$$

where L' is the number of samples in the down-sampled signal and t is the segment size in seconds. In this work, for each down-sampled segment $DS_i(L')$, f_s is 72 Hz, t is 5 seconds, and L' is 360. Using the value of f_s in (1) and the values of t and L' in (2), we can obtain the level of decomposition, $Level = 6$.

In the rest of this paper, we will use the new sampling frequency, i.e., the frequency obtained after down-sampling each segment, which is 72 Hz. According to the Nyquist sampling theorem, the highest frequency component that exists in each down-sampled segment is 36 Hz. Table I shows the wavelet coefficients and the corresponding frequency range. From the table, it can be observed that the approximation coefficients at level 6, labeled as $A_i^\delta(L')$, captures frequencies below 0.67 Hz, as recommended by the AHA. Therefore, approximation coefficients at level 6 is a good estimate of the baseline-wander.

TABLE I. WAVELET COEFFICIENTS AND THE CORRESPONDING FREQUENCY RANGE.

Level	Frequency Range (Hz)
D1	18–36
D2	9–18
D3	4.5–9
D4	2.25–4.5
D5	1.125–2.25
A6	0–0.5625

Next, the simulation is performed for the selection of the optimal mother wavelet function. For each down-sampled segment $DS_i(L')$, we first apply different mother wavelet functions to extract the approximate coefficients at level 6,

i.e., $A_i^\delta(L')$, and then we use third-degree polynomial fitting for each $DS_i(L')$. We assume that the polynomial fitting represents an estimated baseline, which can serve as a standard benchmark for all the mother wavelet functions.

In order to evaluate the performance of the mother wavelet functions to obtain the optimal one, we use the sum of the squared residuals (SSR) and cross-correlation coefficient techniques. We compute SSR between the approximation coefficients $A_i^\delta(L')$ of the mother wavelets and the third-degree polynomial fitting of each down-sampled segment $DS_i(L')$ as follows

$$SSR = \sum_{j=1}^{L'} (y_j - \hat{y}_j)^2. \quad (3)$$

The cross correlation coefficient is evaluated as

$$r_{y\hat{y}} = \frac{\sum_{j=1}^{L'} (y_j - \bar{y}_j)(\hat{y}_j - \bar{\hat{y}}_j)}{\sqrt{\sum_{j=1}^{L'} (y_j - \bar{y}_j)^2 \sum_{j=1}^{L'} (\hat{y}_j - \bar{\hat{y}}_j)^2}}, \quad (4)$$

where y_j is the actual observed value, i.e., the approximation coefficients $A_i^\delta(L')$, \hat{y}_j is the predicted value, i.e., polynomial fit, \bar{y}_j is mean of y_j , and $\bar{\hat{y}}_j$ is the mean of \hat{y}_j . The quantitative comparative analysis between different mother wavelet functions is shown in Fig. 4. In this experiment, we use 23 records of the “100 series” of the MIT-BIH arrhythmia database.

For a better understanding, we divide each value of the SSR and $r_{y\hat{y}}$ by the maximum value of the SSR and $r_{y\hat{y}}$, respectively. It can be inferred from Fig. 4 that the *coif5* mother wavelet function has the minimum SSR and maximum $r_{y\hat{y}}$ as compared to the other wavelet functions. Therefore, the *coif5* wavelet function is an optimal choice for this work.

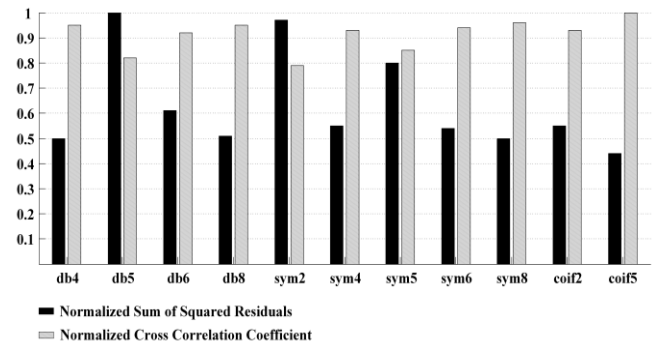


Fig. 4. Comparative analysis of different mother wavelet functions.

The next step is to decompose the signal $DS_i(L')$ into approximate and detail coefficients to level 6 using mother wavelet *coif5*. Of all the coefficients, we will only reconstruct the approximation coefficients at level 6, i.e., $A_i^\delta(L')$, from the wavelet decomposition structure. Fig. 5(a) shows the fifth segment of record 208 taken from the MIT-BIH arrhythmia database. Fig. 5(b) illustrates the estimated

baseline-wander noise using the *coif5* mother wavelet at level 6 on the down-sampled ECG signal. The reconstructed approximation coefficients $A_i^6(L')$ serve as a good estimate for the baseline-wander noise for the down-sampled signal as shown in Fig. 5(b).

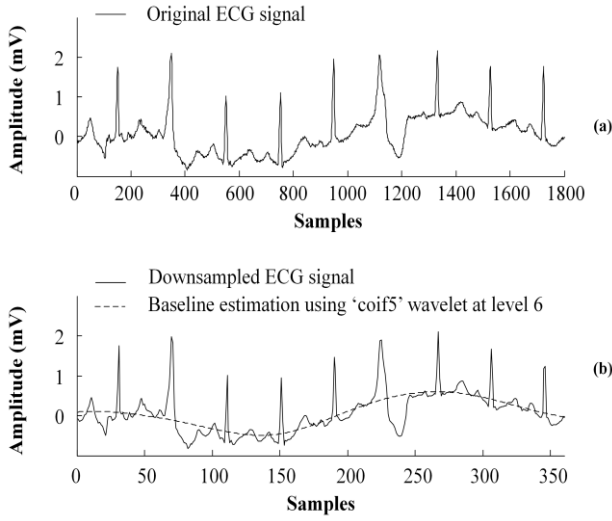


Fig. 5. (a) Original ECG signal taken from the MIT-BIH arrhythmia database record 208 (Signal: MLII), segment no. 5; (b) Down-sampled ECG signal (solid black line) and estimated baseline-wander noise (black dashed line) using the *coif5* wavelet at level 6.

D. Estimation of the Magnitude of Baseline-Wander Noise

For each record of the MIT-BIH arrhythmia database, we first calculate the range R_i of $A_i^6(L')$ for each $DS_i(L')$ using the expression such as

$$R_i = \max(A_i^6(L')) - \min(A_i^6(L')). \quad (5)$$

Next, we perform mean centering of $A_i^6(L')$ by using (6)

$$MCA_i(L') = A_i^6(L') - \text{mean}(A_i^6(L')), \quad (6)$$

where $MCA_i(L')$ is the mean centered approximation coefficients of i th segment, which length is L' . The mean-centering of an approximation coefficients means subtracting the mean value of $A_i^6(L')$ from each data point of $A_i^6(L')$. It results in approximation coefficients with zero mean. Next, we calculate the standard deviation of the mean centered approximation coefficients by using (7)

$$STD_i = \text{std}(MCA_i(L')). \quad (7)$$

In order to estimate the magnitude of the baseline-wander noise, we calculate the cumulative sum as follows

$$CSA_i(L') = \sum_{j=1}^{L'} MCA_i^2(j), \quad (8)$$

where CSA is the cumulative sum of a square of the mean-centered approximation coefficients. Next, we estimate the area under the curve (AUC) of $CSA_i(L')$ using the trapezoidal rule. The basic principle of the trapezoidal rule is to partition the entire segment into multiple sub-intervals, where each sub-interval forms a closed area. On considering

unit spacing for the trapezoidal rule, the AUC is given as

$$AUC_i = \sum_{j=1}^{L'} w_j CSA_i(j), \quad (9)$$

where

$$w_j = \begin{cases} 1/2, & j = 1, \\ 1, & j = 2, \dots, L'-1, \\ 1/2. & j = L'. \end{cases} \quad (10)$$

We have estimated the magnitude of the baseline-wander noise in terms of the AUC. Those segments that have minimal or no baseline-wander noise have a smaller value of AUC.

E. Decision Tree Construction

We use the C4.5 decision tree algorithm to model the classifier using the WEKA data-mining tool. J48 is an open source implementation of the C4.5 decision tree algorithm that is available in WEKA. C4.5 is based on the divide-and-conquer approach, i.e., it selects one attribute of data that most effectively splits the set of samples into subsets. In this work, there are three important attributes that contribute to the likelihood of having large baseline-wander noise. These attributes are range R_i of $A_i^6(L')$, standard deviation STD_i of $MCA_i(L')$, and area under the curve AUC_i of $CSA_i(L')$.

For the MIT-BIH arrhythmia database, we have a total of 17328 segments, i.e., 361 segments for each record. We split the entire database into training and test data sets. The training set consists of 12996 segments for 36 ECG recordings whereas the test data set consists of the remaining 4332 segments for 12 ECG recordings. Using this training data set, a classifier model to predict whether a particular segment of the ECG signal has minimal, moderate, or large baseline-wander noise is built.

Based on the knowledge, regarding baseline-wander noise, the term ‘‘minimal’’ refers to the form of the signal with little to no distortion, ‘‘moderate’’ refers to the form where the baseline-wander displaces some features of the ECG signal, and ‘‘large’’ baseline-wander refers to the form of the signal that is heavily distorted, where baseline-wander displaces the majority of the ECG signal features.

Table II shows the snapshot of the training data set where RID is the row number. The training data set is classified based on extensive experimentation. We assigned weight factors to each attribute according to a certain criteria as mentioned in Table III. For instance, we assigned weight of 1/9 to the range attribute if its value is less than or equal to 0.10. Similarly, we assigned the same weight to AUC and STD attributes if their values are less than or equal to 50 and 0.03, respectively. The remaining weight factors and their corresponding criteria are mentioned in Table III.

After assigning weight factors to each attribute of the RID, we then sum up the weight factors in each row. If the resultant weight is less than or equal to 1/3, then we label the sample as ‘Minimal’. If the resultant weight is greater than 1/3, but less than 2/3, then we label the sample as ‘Moderate’. Similarly, if the resultant weight is greater than

2/3, then the sample is labelled as ‘Large’. For instance, for RID = 1, the value of the weight assigned to attributes range, AUC and STD, are 2/9, 2/9, and 1/9, respectively. The resultant weight is 5/9 and, therefore, we label the sample as ‘Moderate’.

TABLE II. CLASS-LABELLED TRAINING DATA SET.

RID	Range	AUC	STD	Class
1	0.12	86.58	0.03	Moderate
2	0.09	34.70	0.03	Minimal
3	0.07	33.76	0.02	Minimal
4	0.05	6.52	0.01	Minimal
5	0.16	167.87	0.05	Large
:	:	:	:	:
12996	0.14	39.26	0.03	Moderate

TABLE III. WEIGHT FACTORS ASSIGNED TO THE ATTRIBUTES.

Weight Factor	Range	AUC	STD
1/9	≤ 0.1	≤ 50	≤ 0.03
2/9	$> 0.1 \ \& \ \leq 0.25$	$> 50 \ \& \ \leq 100$	$> 0.03 \ \& \ \leq 0.05$
3/9	> 0.25	> 100	> 0.05

In the training data set, we have three attributes. The criteria for the parent split attribute selection are based on the gain ratio. In order to find the gain ratio, we first calculate the residual uncertainty, i.e., entropy in data set D using (10)

$$Entropy(D) = - \sum_{i=1}^n p_i \log_2(p_i), \quad (10)$$

where n is the set of desired classes and p_i is the probability of each class. Next, we calculate the information required to split the data set D using attribute A into j parts by using expression such as

$$Info_A(D) = \sum_{j \in \text{values}(A)} \frac{|D_j|}{|D|} Entropy(D_j), \quad (11)$$

where $\text{values}(A)$ is the set of all possible values for A and D_j is the subset of D , for which attribute A has a value j . For each attribute, we calculate the information gain as follows

$$Gain(D, A) = Entropy(D) - Info_A(D). \quad (12)$$

The information gain ratio of attribute A is given as

$$Gain Ratio_A = \frac{Gain(D, A)}{Split Information(D, A)}, \quad (13)$$

where $Split Information(D, A)$ is calculated as in (14)

$$Split Information(D, A) = - \sum_{k \in C} \frac{|D_k|}{|D|} \log_2 \frac{|D_k|}{|D|}, \quad (14)$$

where D_k is the subset of D having the outcome k of the class C for the attribute A . $Entropy(D)$ is calculated only once, whereas $Info_A(D)$, $Gain(D, A)$, and $Split Information(D, A)$ is calculated for each attribute.

Table IV shows the gain ratio for the range STD and AUC for all the training data. As shown in the table, AUC

provides a higher gain ratio than other attributes, and thus AUC is going to be the parent split attribute. Now, that the decision tree is constructed, we can use this classifier model to predict whether the new test segments have minimal, moderate, or large baseline-wander noise.

TABLE IV. GAIN RATIO ATTRIBUTE EVALUATION.

Parameters	Attributes		
	Range	STD	AUC
Gain(D,A)	0.809	0.845	1.008
Gain Ratio _A	0.233	0.407	0.574
Entropy(D)	1.214		

IV. RESULTS

In this section, the proposed methodology is tested on real ECG signals taken from the MIT-BIH arrhythmia database. The database contains 48 ECG records recorded at a sampling frequency of 360 Hz and 11-bit resolution over a 10-mV range [29], [30]. Each record in the MIT-BIH arrhythmia database is slightly longer than 30 min. We partition each record into 5 second, non-overlapping segments, which results in 361 segments for each record.

The proposed methodology is illustrated in Fig. 6. In order to give a better understanding about the proposed method, we have considered segment no. 5, i.e., 0:20 to 0:25 (min:s) of mitdb/205 and mitdb/208 from the MIT-BIH arrhythmia database as shown in Fig. 6(a) and Fig. 6(e), respectively. The original ECG signal is down-sampled by 5. The down-sampled data (solid line) in Fig. 6(b) and Fig. 6(f) is then decomposed into approximate and detail coefficients up to level 6 by using the mother wavelet *coif5*. The approximation coefficients (dashed line) at level 6 serve as a good estimate for the baseline of the down-sampled data. At the next step, mean-centering of the approximation coefficients is performed, i.e., the mean of approximation coefficients is eliminated from each data point. The dashed-dotted line represents the mean-centered approximation coefficients in Fig. 6(b) and Fig. 6(f).

The mean-centered approximation coefficients are squared as shown in Fig. 6(c) and Fig. 6(g). We compute the cumulative sum of the squared mean-centered approximation coefficients and then calculate the area under the curve using the trapezoidal rule to estimate the magnitude of the baseline-wander noise. As shown in Fig. 6(d) and Fig. 6(h), segment no. 5 of mitdb/205 has an area of 28.35, whereas mitdb/208 has an area of 6794.68. From the figure, it can be inferred that minimal baseline-wander noise results in a smaller area, whereas large baseline-wander noise results in a larger area.

For the purpose of demonstration, two records (record 123 and record 201) from the MIT-BIH arrhythmia database are considered. For each segment, we reconstruct the approximation coefficients of the down-sampled data using the *coif5* wavelet at level 6, and then calculate the range of approximation coefficients. Fig. 7(a) and Fig. 7(c) show the range for each segment of mitdb/123 and mitdb/201, respectively. Next, we estimate the area under the curve for the cumulative sum of the squared mean-centered approximation coefficients, as shown in Fig. 7(b) and Fig. 7(d), for each segment of mitdb/123 and mitdb/201, respectively.

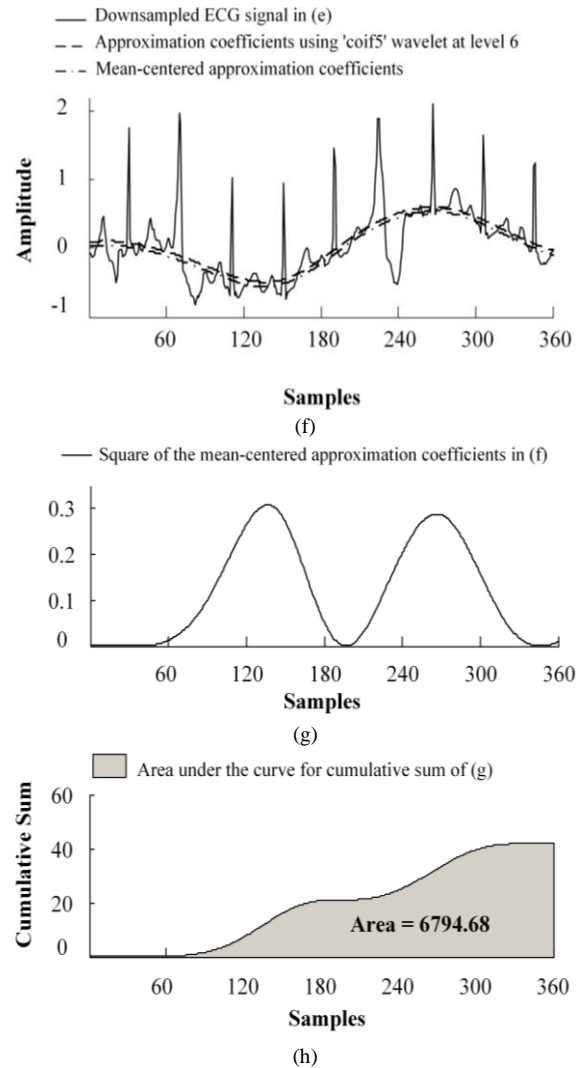
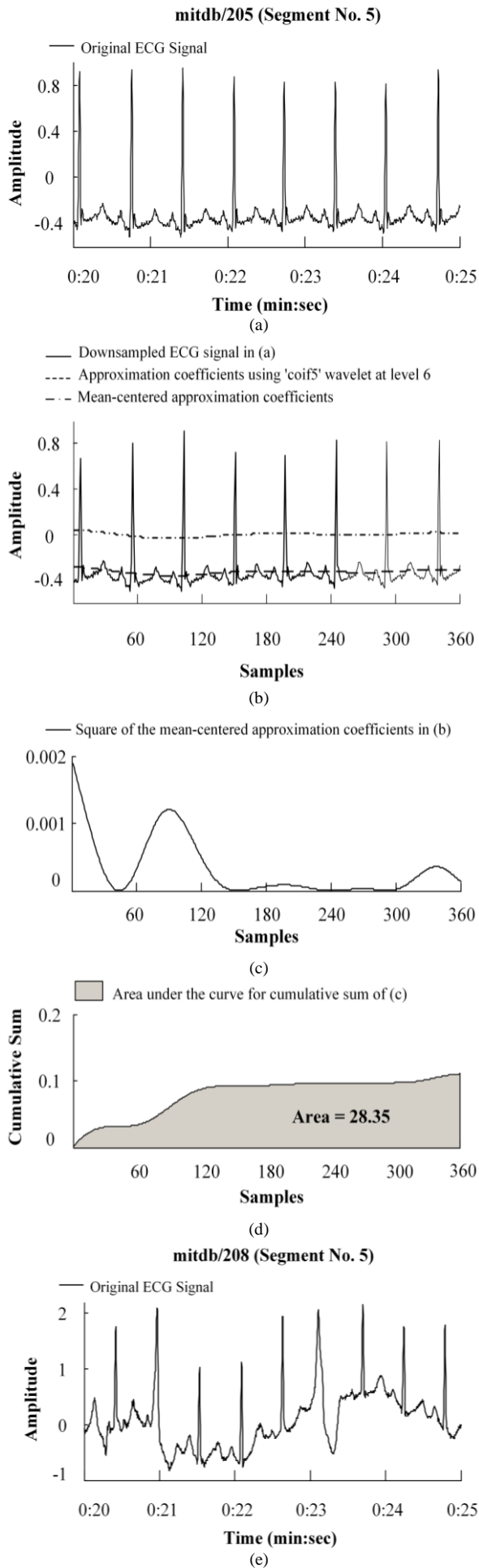


Fig. 6. (a) Original ECG signal taken from mitdb/205 (Segment no. 5); (b) Mean-centering of approximation coefficients on the down-sampled ECG signal in (a). The solid line denotes the down-sampled ECG signal, the dashed line denotes the approximation coefficients that were obtained using the *coif5* wavelet at level 6, and the dashed-dotted line denotes the mean-centered approximation coefficients; (c) Square of the mean-centered approximation coefficients in (b); (d) Area under the curve for the cumulative sum of mean-centered approximation coefficients in (c) using the trapezoidal rule; (e) Original ECG signal taken from mitdb/208 (Segment no. 5); (f) Mean-centering (dashed-dotted line) of approximation coefficients (dashed line) on the down-sampled (solid line) ECG signal in (e); (g) Square of the mean-centered approximation coefficients in (f); (h) Area under the curve for the cumulative sum of the mean-centered approximation coefficients in (g).

For better visualization, the AUC from segment no. 60 to segment no. 180 is zoomed. From Fig. 7, it can be observed that some segments have a very large baseline-wander while the rest have minimal or no baseline. In addition, segment nos. 22 and 221 have the maximum baseline-wander in mitdb/123 and mitdb/201, respectively.

In order to classify each segment of every record as having minimal, moderate, or large baseline-wander, a classifier using the decision tree algorithm is modelled. Fig. 8 shows the baseline-wander decision tree classifier based on the training data set, the snapshot of which is provided in Table II.

Now, that the decision tree is constructed, we can use it as a basis for predicting whether the new segment has a minimal, moderate, or large baseline-wander noise. The

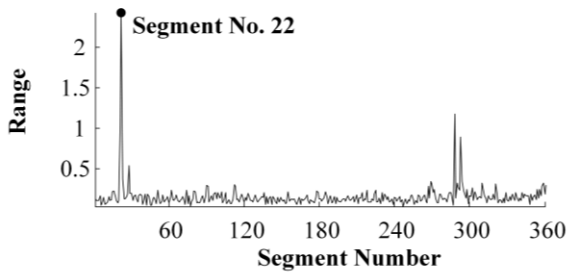
remaining 4332 segments of 12 ECG recordings as the test data are used.

We consider two different sequences of segments, mentioned in Table V, to demonstrate the prediction obtained with the test data. The first sequence consists of segment nos. 236, 237, and 238 of mitdb/108, whereas the second sequence consists of segment nos. 287, 288, and 289 of mitdb/123 from the MIT-BIH arrhythmia database.

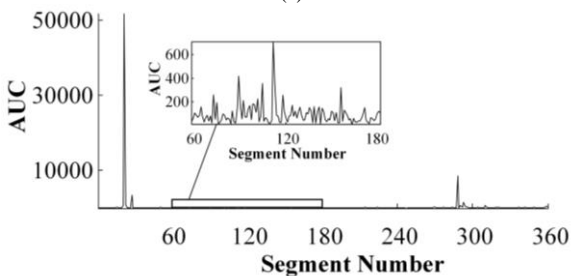
TABLE V. ATTRIBUTES OF SIX NEW TUPLES FOR DEMONSTRATING THE CLASSIFIER MODEL.

Record no.	Segment no.	Attributes		
		Range	STD	AUC
108	236	0.042	0.01	5.41
	237	0.163	0.05	144.96
	238	0.732	0.24	1976.70
123	287	0.048	0.01	7.95
	288	1.179	0.38	8507.10
	289	0.139	0.04	71.29

mitdb/123

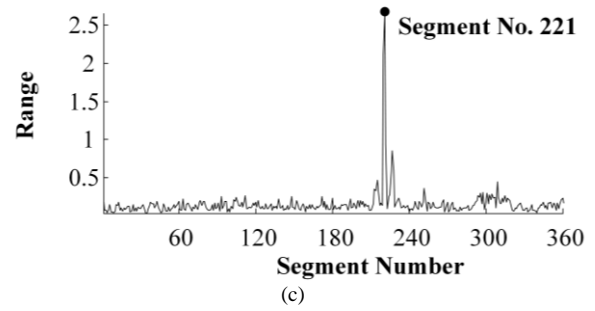


(a)

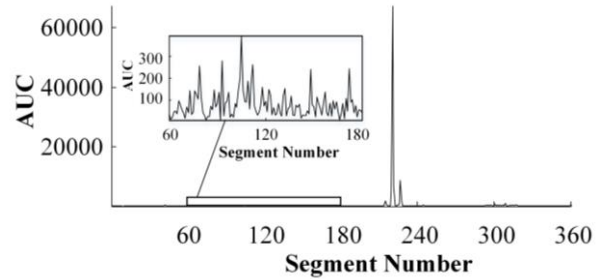


(b)

mitdb/201



(c)



(d)

Fig. 7. For each segment of mitdb/123 from the MIT-BIH arrhythmia database, (a) range of approximation coefficients and (b) the area under the curve for the cumulative sum of the squared mean-centered approximation coefficients. For each segment of mitdb/201 from the MIT-BIH arrhythmia database, (c) range of approximation coefficients and (d) is the area under the curve for the cumulative sum of the squared mean-centered approximation coefficients.

According to the decision tree of Fig. 8, we first check the value of AUC. As for the first new tuple, i.e., segment no. 287 of record no. 123, the value of AUC is 7.95, we branch to the node labeled RANGE. Next, as the RANGE of the newly added tuple is less than 0.09, we branch to the node labeled AUC. The decision tree again checks for AUC, which in our case is less than 49.96, we branch to the node labeled “MINIMAL.” In other words, we predict, with high probability that the new segment has minimal baseline-wander noise. The same steps are repeated for all the new tuples.

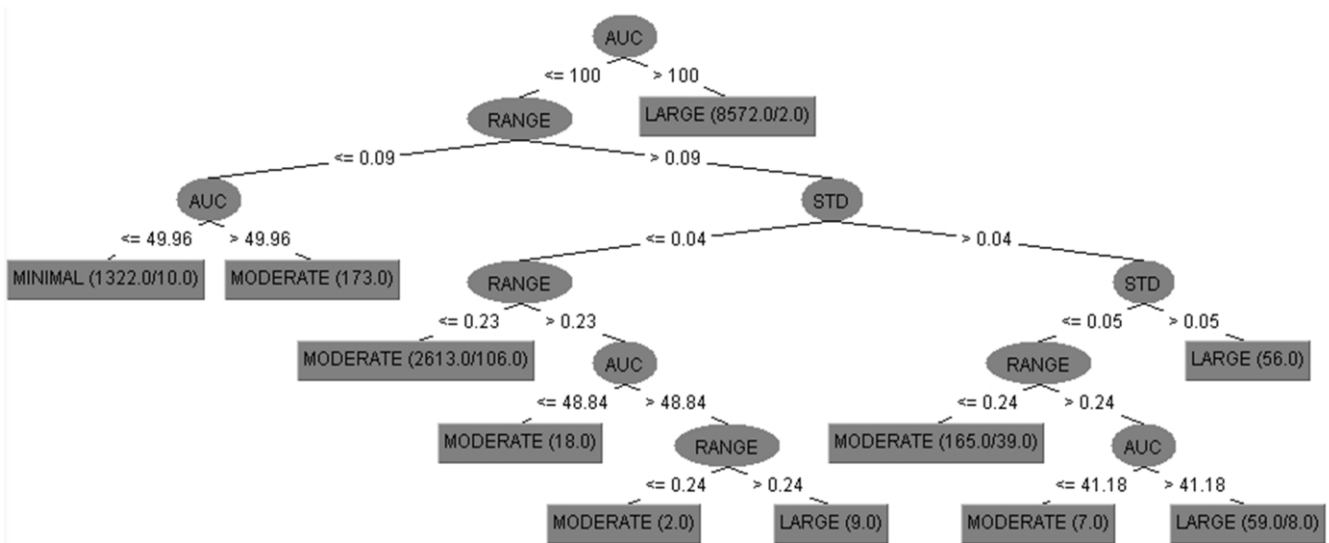


Fig. 8. Baseline-wander decision tree based on the training data set provided in Table II.

Fig. 9 illustrate the prediction of the classifier model for the six tuples listed in Table V as test data segments. It can

be observed that the classifier model can successfully predict the class of each new tuple even for two successive segments

that have a sharp transition from minimal to large baseline-wander as shown in Fig. 9.

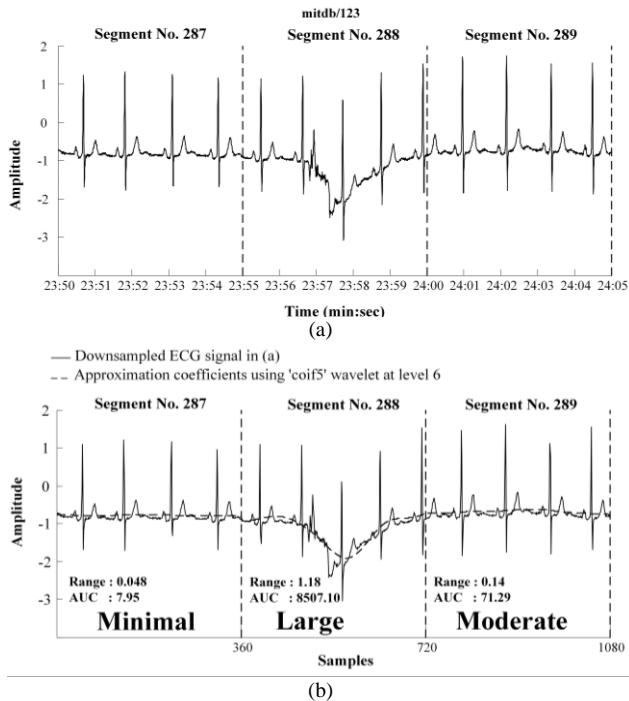


Fig. 9. Segment nos. 287, 288, and 289 of mitdb/123 as the test data: (a) Original ECG signal; (b) prediction result of the classifier model for each segment.

Table VI shows the confusion matrix of the model classifier for the test data. In the below table, the diagonal elements represent correctly classified instances, whereas off-diagonal elements represent the misclassified instances. As shown in the table, there are a total of 4332 instances, out of which 48 instances are incorrectly classified.

TABLE VI. CONFUSION MATRIX OF THE MODEL CLASSIFIER FOR THE TEST DATA.

	PREDICTED		
	Minimal	Moderate	Large
Minimal	454	35	0
Moderate	2	928	2
Large	0	9	2902

Table VII shows the class-wise performance of the model classifier. The proposed method can identify minimal baseline-wander segments in ECG signal with the accuracy of 99.15%. The obtained results show that the model classifier can dependably classify baseline-wander noise in ECG signals with an overall accuracy of 98.89%.

TABLE VII. CLASS-WISE PERFORMANCE OF THE MODEL CLASSIFIER ON TEST DATA.

	CLASS		
	Minimal	Moderate	Large
True Positive	454	928	2902
True Negative	3841	3356	1419
False Positive	2	44	2
False Negative	35	4	9
Sensitivity (%)	92.84	99.57	99.69
Precision (%)	99.56	95.47	99.93
Specificity (%)	99.95	98.70	99.86
False Pos. Rate (%)	0.052	1.29	0.141
False Neg. Rate (%)	7.16	0.43	0.31
Accuracy (%)	99.15	98.90	99.74
F-measure (%)	96.1	97.5	99.8

V. DISCUSSION

Baseline wandering is a well-known classical problem. High pass digital filtering is the most common method to eliminate the baseline-wander. FIR filters introduce ringing effect at the start of the ECG signal. In addition, FIR filters are computationally intensive and introduce long time delay. On the contrary, IIR filters have non-linear phase response that can be resolved by applying bidirectional filtering. This technique is implemented for off-line ECG signal processing. However, if some time delay is accepted, then IIR filters can be used for on-line processing. One of the disadvantages of digital filtering is fixed cut-off frequency. If the cut-off frequency is set too low, i.e., 0.05 Hz, the filtering may not significantly suppress the baseline-wander. However, if the cut-off frequency is set to the AHA recommended frequency, i.e., 0.67 Hz, the filtering may introduce distortion in the baseline-free segments of ECG signal. Moreover, the spectral contents of the baseline-wander signal may contain frequencies higher than 0.67 Hz. Therefore, setting a fixed cut-off frequency for digital filtering is not an optimal choice.

Another approach is polynomial interpolation. The accuracy of the interpolation methods highly depends on the accurate determination of knots. The method fails if the knots are not accurately identified. Cubic spline interpolation is not recommended in the presence of sudden baseline shifts.

Lenis, *et al.* [31] compares baseline-wander removal techniques and concludes that wavelet-based baseline cancellation outperforms other methods in terms of preserving the ST-segment of ECG signal. However, it is computationally intensive as compared to Butterworth filtering or polynomial interpolation.

Blanco-Velasco, *et al.* [25] proposes EMD based method to remove the baseline-wander from ECG signal. One of the major drawbacks of EMD method is mode mixing effect. Also, EMD method is computationally very demanding.

Shusterman, *et al.* [11] proposes a two-step method for selective filtering of ECG signal by combining interpolation and filtering techniques. In the first step, the magnitude of the baseline-wander is determined and ECG signal is classified as small or large. If large baseline-wander is present, the segments are filtered using a bidirectional high-pass filter. In the second step, small baseline-wander or any residual left during the first step is eliminated by linear interpolation. The approach is tested on 100 simulated, 110 ambulatory, and 100 segments obtained from the 32-channel body surface potential mapping ECGs. This approach is computationally intensive as it requires bidirectional filtering, frequency estimation, and polynomial fitting for each 20 seconds of data. Moreover, this approach may not be effective for the correction of baseline-wander, which has a frequency higher than 0.675 Hz.

In this paper, a method for classifying baseline-wander noise in ECG signals is developed. The proposed method works on the down-sampled data of each segment to reduce the data bandwidth as well as computational load. Moreover, the down sampling factor is selected such that it does not require a bidirectional antialiasing filtering. The magnitude

of the baseline-wander noise is estimated by calculating AUC for the cumulative sum of squared mean-centered approximation coefficients. The results indicate that those segments that have minimal baseline-wander have the smaller value of AUC. The C4.5 decision tree algorithm to model the classifier is used. The results indicate that the model classifier achieves an accuracy of 99.15 % in identifying minimal or no baseline-wander segments.

Present study introduces a new class, labeled as moderate baseline-wander, for the first time. These are the segments where baseline-wander displaces some features of the ECG signal. These segments can neither be classified as minimal nor large baseline-wander.

VI. CONCLUSIONS

In this study, we present a novel method for the classification of baseline-wander noise in ECG signals. Each segment of the ECG signal is classified as minimal, moderate, or large baseline-wander noise. The results indicate that the model classifier achieves an accuracy of 99.15 %, 98.90 %, and 99.74 % in identifying minimal, moderate, and large baseline-wander segments, respectively.

In recent years, many researchers have proposed baseline-wander removal techniques. Each of these techniques has advantages and drawbacks in terms of accuracy and computational cost. Baseline-wander removal should be considered as an optimization problem. The results show that the proposed method can possibly be used in devising an algorithm to optimize the baseline-wander removal in terms of computational cost and accuracy. The devised algorithm should be based on selective filtering of moderate and large baseline-wander segments in ECG signals by combining different baseline-wander removal techniques to achieve the best trade-off between accuracy and computational cost.

In the future, we would like to investigate the application of proposed method on other types of noise in ECG signal.

REFERENCES

- [1] A. A. R. Bsoul, S. Ji, K. Ward, and K. Najarian, "Detection of P, QRS, and T components of ECG using wavelet transformation", in *2009 ICME International Conference on Complex Medical Engineering*, pp. 1-6. DOI: 10.1109/ICME.2009.4906677.
- [2] J. A. G. Gnecci, A. d. J. V. Herrejón, A. d. C. T. Anguiano, A. M. Patiño, and D. L. Espinoza, "Advances in the construction of ECG wearable sensor technology: The ECG-ITM-05 eHealth data acquisition system", in *Proc. 2012 IEEE Ninth Electronics, Robotics and Automotive Mechanics Conference*, pp. 338-342. DOI: 10.1109/CERMA.2012.60.
- [3] A. Dosinas, R. Lukocius, M. Vaitkunas, G. Nedzinskaite, P. Vaskys, S. Gudzius, and A. Jonaitis, "Sensors and signal processing methods for a wearable physiological parameters monitoring system", *Elektronika ir Elektrotechnika*, vol. 23, no. 5, pp. 74-81, 2017. DOI: 10.5755/j01.eic.23.5.19245.
- [4] Y-W. Bai, C-Y. Cheng, C-K. Lu, C-H. Huang, Y-T. Chen, and Y-N, "Adjustable 60 Hz noise reduction and ECG signal amplification of a remote electrocardiogram system", in *Proc. of the 20th IEEE Instrum. Technol. Conf.*, 2003, pp. 197-202. DOI: 10.1109/IMTC.2003.1208151.
- [5] L. Fay, V. Misra, and R. Sarpeshkar, "A Micropower electrocardiogram amplifier", *IEEE Trans. on Biomed. Circ. and Sys.*, vol. 3, no. 5, pp. 312-320, 2009. DOI: 10.1109/TBCAS.2009.2026483.
- [6] C. So-In, C. Phaudphut, and K. Rujirakul, "Real-time ECG noise reduction with QRS complex detection for mobile health services", *Arab J. Sci. Eng.*, vol. 40, no. 9, pp. 2503-2514, 2015. DOI: 10.1007/s13369-015-1658-1.
- [7] F. A. Afsar, M. S. Riaz, and M. Arif, "A comparison of baseline removal algorithms for electrocardiogram (ECG) based automated diagnosis of coronary heart disease", in *Proc. 3rd Int. Conf. Bioinform. Biomed. Eng.*, pp. 1-4. DOI: 10.1109/ICBBE.2009.5162156.
- [8] X. Hu, Z. Xiao, and N. Zhang, "Removal of baseline wander from ECG signal based on a statistical weighted moving average filter", *J. of Zhejiang Univ. SCIENCE C*, vol. 12, no. 5, pp. 397-403, 2011. DOI: 10.1631/jzus.C1010311.
- [9] N. Sonal and R. Shastri, "Comparative analysis of ECG baseline wander correction techniques", in *2014 Int. Conf. Commun. Signal. Proc.*, pp. 762-766.
- [10] R. Lakhwani, S. Ayub, and J. P. Saini, "Design and comparison of digital filters for removal of baseline wandering from ECG signal", in *Proc. 5th Int. Conf. Comput. Intell. Comm. Netw.*, 2013, pp. 186-191. DOI: 10.1109/CICN.2013.48.
- [11] V. Shusterman, S.I. Shah, A. Beigel, and K.P. Anderson, "Enhancing the precision of ECG baseline correction: Selective filtering and removal of residual error", *Comput. Biomed. Res.*, vol. 33, no. 2, pp. 144-160, 2000. DOI: 10.1006/cbmr.2000.1539.
- [12] J. Kozumplik and I. Provaznik, "Fast time-varying linear filters for suppression of baseline drift in electrocardiographic signals", *Biomed. Eng.*, pp. 1-16, 2017. DOI: 10.1186/s12938-017-0316-0.
- [13] S. A. A. Sheikh, N. Razzaq, and T. Zaidi, "Baseline wander removal from ECG signal using state space recursive least squares (SSRLS) adaptive filter", in *Proc. 2nd Int. Conf. Robot Artif. Intell. (ICRAI)*, 2016, pp. 58-62. DOI: 10.1109/ICRAI.2016.7791229.
- [14] A. A. Hiasat, M. M. Al-Ibrahim, and K. M. Gharai, "Design and implementation of a new efficient median filtering algorithm", *IEEE Proc. Vision, Image Signal Proc.*, vol. 146, no. 5, pp. 273-278, 1999. DOI: 10.1049/ip-vis:19990444.
- [15] V. K. Pandey, "Adaptive filtering for baseline wander removal in ECG", in *Proc. 10th IEEE Int. Conf. Inf. Technol. Appl. Biomed.*, 2010, pp. 1-4. DOI: 10.1109/ITAB.2010.5687642.
- [16] K. L. Park, K. J. Lee, and H. R. Yoon, "Application of a wavelet adaptive filter to minimize distortion of the ST-segment", *Med. Biol. Eng. Comput.*, vol. 36, no. 5, pp. 581-586, 1998. DOI: 10.1007/BF02524427.
- [17] H. He, Y. Tan, and Y. Wang, "Optimal base wavelet selection for ECG noise reduction using a comprehensive entropy criterion", *Entropy*, vol. 17, no. 9, pp. 6093-6109, 2015. DOI: 10.3390/e17096093.
- [18] E. Castillo, D. P. Morales, A. García, L. Parrilla, N. Lopez-Ruiz, A. J. Palma, "One-step wavelet-based processing for wandering and noise removing in ECG signals", in *Proc. Int. Work-Confer. Bioinf. Biomed. Eng.*, 2013, pp. 491-498. DOI: 10.1155/2013/763903.
- [19] S. K. Yadav, R. Sinha, and P. K. Bora, "Electrocardiogram signal denoising using non-local wavelet transform domain filtering", *IET Signal Proc.*, vol. 9, no. 1, pp. 88-96, 2015. DOI: 10.1049/iet-spr.2014.0005.
- [20] V. Pockevicius, M. Cepenas, D. Miklusis, V. Markevicius, L. Zabulienė, D. Navikas, A. Valinevicius, D. Andriukaitis, "Feasibility research of non-invasive methods for interstitial fluid level measurement", *Bio-Medical Materials and Engineering*, vol. 28, iss. 2, pp. 601-612. 2017. DOI: 10.3233/BME-171699.
- [21] O. Guven, A. Eftekhari, W. Kindt, and T.G. Constandinou, "Computationally efficient real time interpolation algorithm for non-uniform sampled bio signals", *Healthc. Technol. Lett.*, vol. 3, no. 2, pp. 105-110. DOI: 10.1049/htl.2015.0031.
- [22] U. Satiya, B. Ramkumar, and M. S. Manikandan, "A robust sparse signal decomposition framework for baseline wander removal from ECG signal", in *Proc. IEEE Reg. 10th Conf. (TENCON)*, 2016, pp. 2470-2473. DOI: 10.1109/TENCON.2016.7848477.
- [23] R. Silipo and C. Marchesi, "Artificial neural networks for automatic ECG analysis", *IEEE Trans Signal Proc.*, vol. 46, no. 5, pp. 1417-1425, 1998. DOI: 10.1109/78.668803.
- [24] J. Kuzilek, V. Kremen, F. Soucek, and L. Lhotska, "Independent component analysis and decision trees for ECG Holter recording denoising", *PLoS ONE*, vol. 9, no. 6, pp. 1-9, 2014. DOI: 10.1371/journal.pone.0098450.
- [25] B. Weng, M. Blanco-Velasco, and K. E. Barner, "ECG signal denoising and baseline wander correction based on the empirical mode decomposition", *Comp. Biol. Med.*, vol. 38, no. 1, pp. 1-13, 2008. DOI: 10.1016/j.compbiomed.2007.06.003.
- [26] T. Tanantong, "A study on the effects of window size on electrocardiogram signal quality classification", in *Proc. Knowl.*

- Manag. Int. Conf. (KMICe)*, 2016, pp. 333–338.
- [27] VK. Murthy, TM. Grove, GA. Harvey, LJ. Haywood, “Clinical usefulness of ECG frequency spectrum analysis”, in 1978 *Proc. Annual Symp. Comp. Appl. Med Care*, pp. 610-612.
- [28] N. W Willigenburg, A. Daffertshofer, I. Kingma, and J.H.V. Dieën, “Removing ECG contamination from EMG recordings: A comparison of ICA-based and other filtering procedures”, *J. Electrom. Kinesiol.*, vol. 22, no. 3, pp. 485–493. DOI: 10.1016/j.jelekin.2012.01.001.
- [29] The MIT-BIH Arrhythmia Database. [Online]: <http://physionet.org/physiobank/database/mitdb/>
- [30] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, *et al.*, “PhysioBank, physioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals”, *Circulation*, vol. 101, no. 23, pp. 215–220, 2000. DOI: 10.1161/01.cir.101.23.e215.
- [31] G. Lenis, N. Pilia, A. Loewem W. H. W. Schulze, and O. Dössel, “Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG: A simulation study”, *Comput. Mathem. Meth. Med.*, pp. 1–13, 2017. DOI: 10.1155/2017/9295029.