2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)
4-7 March 2018
Las Vegas, Nevada, USA

A Predictive Framework for ECG Signal Processing Using Controlled Nonlinear Transformation

Jiaming Chen[1] and Abolfazl Razi[2]

***Abstract*— In this paper, a novel method is proposed to predict upcoming heart abnormalities by processing electrocardiogram (ECG) signals. The core idea behind the proposed method is to use a controlled nonlinear transformation to project extracted signal features into a higher-order dimensional space with desired geometric properties. In particular, we enforce the projected clusters of different abnormalities to symmetrically encircle the normal cluster through penalizing the clustering non-symmetry. An immediate utility of this method is to characterize the deviation of ECG signal samples from the patient-specific norms towards different abnormality classes. Moreover, this method can be used to enhance our prediction about the potential upcoming heart problems before their occurrences. This is a critical point, since a timely diagnosis and therapeutic intervention can significantly reduce the heart-related mortality rate. We applied this method to publicly available MIT-BIH dataset with 3 abnormality classes and the results suggest, respectively, 8%, 9% and 12% improvement in predicting the three abnormality classes. The proposed framework is general and applicable to a broad range of biomedical signals.**

I. INTRODUCTION

In today's health care industry, computer science techniques are frequently used to assist physicians in making clinical diagnoses. As an important application of biomedical signal processing, automated classification of electrocardiogram (ECG) signals has been investigated by researchers in the past decade [1]. Since abnormal cardiac cycles can indicate potential life-threatening diseases (e.g. arrhythmia), such systems are often used in intensive care unit (ICU), wearable monitor kit, etc. to provide continuous and real-time monitoring.

The common spirit of heart monitoring methods is using large datasets of annotated ECG signals to model the impact of various heart disorders on the ECG signal morphology in order to recognize heart disorders for test signals. Several conventional classification algorithms, such as fuzzy C-means clustering (FCM) [2], support vector machine (SVM) [3], artificial neural network (ANN) [4] are reported for ECG signal classification.

One of the issues in processing ECG signals is the intrinsic morphological variability of ECG waveforms even among healthy individuals due to gender, age, race, environmental conditions, as well as patient-specific anatomy, genetics and metabolism [1], [5]. In order to accommodate these natural variations, several patient-specific classification methods are proposed using mixture of experts [5], and block-based neural network [6].

An important remaining issue for all of the aforementioned methods is presenting the results in terms of normal or abnormal classes without characterizing the severity of the abnormality. Further, these methods are often tuned to avoid reporting false alarms [7]; thus may miss minor abnormalities which are worthy of paying a closer attention [8].

In this paper, we propose a novel prediction framework using a robust deviation analysis. More specifically, a training dataset of ECG signals is processed to recognize the geometry of normal and various abnormality classes in the feature space. Subsequently, a controlled nonlinear transformation is developed in order to project data into a higher order dimensional space, such that different abnormality clusters symmetrically surround the normal cluster. The patient-adaptable coefficients for this nonlinear transformation are tuned using a multi-objective particle swarm optimization (MOPSO) [9], where two objective functions balance between the clustering *cohesion* and *symmetry*. With this approach, in addition to reporting major alarms similar to conventional methods, we report any significant deviations of normal samples towards any of abnormality cases as minor alarms for further analysis. This approach can significantly reduce heart-related mortality rate by helping a timely diagnosis and prediction of upcoming severe heart abnormalities.

II. ECG DATA PROCESSING

In this work, the benchmark database, MIT-BIH arrhythmia database (MITDB) [10] is used to evaluate the performance of the proposed method. This database includes a total of 48 ECG signal recordings with a uniform sampling frequency of 360Hz. According to the American association of medical instrumentation (AAMI)'s recommendations [11], we divide the 44 tapes into two non-overlapping subsets (DS1 and DS2), each including 22 tapes and approximately 50,000 cardiac cycles. The cardiac cycles include five classes, namely N (normal and bundle branch block beats), V (ventricular ectopic beats), S (supraventricular ectopic beats), and F (fusion of N and V beats).

A. *Smoothing and denoising*

The Daubechies 8 order wavelet is used for ECG signal denoising, since it is proven to be more efficient than other basis functions [12]. If an ECG signal with a different sampling frequency is used, we can find the number of decomposition level L to obtain the same level of smoothness as follows:

$$L = 1 + \lfloor \log_2 \frac{Fs}{360} \rfloor, \quad (1)$$

[1] Chen Jiaming is a Masters Student in the School of Electrical Engineering at Southwest Jiaotong University, Chengdu 610031, China, and also in the School of Informatics, Computing and Cyber Systems at Northern Arizona University, Flagstaff, Arizona 86001, Email: jiaming_chen@nau.edu

[2] Abolfazl Razi is an Assistant Professor of Informatics, Computing and Cyber Systems at Northern Arizona University, Email: abolfazl.razi@nau.edu.

978-1-5386-2405-0/18/$31.00 ©2018 IEEE

TABLE I: Training and test datasets in MITDB.

	Number of segments per AAMI class				
Evaluation Dataset	N	V	S	F	Total
DS1:Training	12633	2053	550	121	15357
DS2:Test	11721	2356	862	256	15195
Total	24354	4409	1412	377	30552

where, the floor function $\lfloor x \rfloor$ represents the largest integer lower than x. By discarding the detail coefficients at the first level, D_1, we eliminate the electrosurgical and muscle noise which is typically within the 100khz-10MHz range. Furthermore, a five ordered fitted polynomial is subtracted from the ECG signal to remove low frequency noise terms.

B. *Segmentation*

A cardiac cycle is represented by its five fiducial peaks: P, Q, R, S and T. The accurate detection of QRS complex, the most dominant peak within one cycle, is critical for the subsequent analysis. Cardiac cycles typically exhibit QRS complexes in the frequency range of 5 to 22Hz [13]. The detail coefficients of level 4 (D_4) and level 5 (D_5) include the information of QRS complexes. Based on maximal overlap discrete wavelet transform (MODWT), the algorithm proposed in [14], is applied to detect R peaks.

To delineate ECG waveform by eliminating transient terms, here we combine three consecutive cycles to form a representative segment, following the work in [8]. A segment is labeled Normal (N), if all member cycles are annotated as N, otherwise it is labeled as the only abnormal class (V, S or F) within the segment. If a segment includes two or more different abnormal cycles, the segment is discarded to avoid fake transient events. In the subsequent analysis, one segment is regarded as one data sample. Table.I summarizes the sample types in DS1 and DS2.

C. *Feature Extraction*

Each cardiac cycle can be represented as a set of features. As discussed in [1], [11], [13], [15], a collective analysis of features in three categories, namely temporal, morphological and frequency domain, reflects the main differences between arrhythmia types. Table.II summarizes the extracted features. Here, cycle-based features (SET 1) include the mean and standard deviation of the corresponding features across three cycles within a segment, while segment-based features (SET 2) are calculated once per segment. For each segment, a set of 22 features are extracted and normalized to yield zero mean unit variance features. We represent the k^{th} cardiac segment with start time t_k and duration τ_k as a 22×1 vector $\mathbf{x}_k$:

$$g : s(t_k : t_k + \tau_k) \mapsto \mathbf{x}_k, \tag{2}$$

where $g()$ represents the segmentation and feature extraction.

III. Personalized Classification Method

The proposed framework processes the pre-processed data samples $\mathbf{x}_k$ and provides *major (red)* and *minor (yellow) alarms* based on the severity of the deviation of the current data sample from the patient-specific normal trend. A *red alarm* means that the current sample does not pass the initial test using a global classifier. Samples classified as normal go through a personalized investigation. If the deviation of normal samples from the normal cluster centroid exceeds a predefined threshold, we set a *yellow alarm* of the corresponding type. This process involves the following stages.

TABLE II: Features extracted from ECG signal

Feature Type	SET1	SET2
Temporal Features	QRS duration, QT duration, PR duration	mean$(R_{i+1}-R_i)$, mean$(R_i - R_{avg})$
Morphological Features	max positive peak to second peak ratio	signal average energy, max positive peak, max negative peak, peak to energy ratio
Frequency Domain Features	signal power level at 7.5Hz, 10Hz, 12.5Hz, 15Hz	

A. *Global Classifier*

Each segment $\mathbf{x}_k \in \Omega^d$ in the original feature space of dimension d associates with a true label $y_k \in \mathcal{Y}$. In general, $\Omega^d = \{\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_C\}$ includes one normal class $\mathcal{X}_1$ and $C-1$ abnormal classes $\mathcal{X}_2, \ldots, \mathcal{X}_C$. In this work, we have $C = 4$, $d = 22$, $\Omega^{22} = \{\mathcal{N}, \mathcal{S}, \mathcal{V}, \mathcal{F}\}$ and $\mathcal{Y} = \{N, V, S, F\}$. The goal of classification is to design a function $\hat{y}_k = f(\mathbf{x}_k)$, such that the pairwise distances between the true and predicted labels for training samples $\sum_{k=1}^{N} d(\hat{y}_k, y_k)/N$ are minimized based on desired distance metric.

A global classifier denoted by $f_1()$ is trained using data samples $\mathbf{x}_k$ in the training set (DS1) according to the guidelines offered by AAMI [16]. This classifier $f_1 : \mathbf{x}_k \mapsto \hat{y}_k^1$ maps each sample into one of the 4 classes in $\mathcal{Y}$. If $\hat{y}_k^1 \neq N$, a red alarm of type $\hat{y}_k^1$ is fired, otherwise $\mathbf{x}_k$ is delivered to the subsequent deviation analysis. We arbitrarily choose k-nearest neighbors (kNN) with $k = 10$ for its low complexity and high accuracy. However, the proposed framework is general and does not depend on the choice of classifier.

B. *Deviation Analysis Moduler*

Here, we quantify the deviation of samples from the normal cluster centroid towards any of the abnormal classes. This step is initialized after collecting the set of the first 20% of segments labeled normal for user i, denoted as $\mathcal{N}_0^i = \{\mathbf{x}_k(i) : k = 1, 2, \ldots, \lfloor n_i/5 \rfloor, y_k(i) = N\}$. Likewise, the global abnormal clusters ($\mathcal{V}$, $\mathcal{S}$, $\mathcal{F}$) consist of abnormal samples in DS1. Now, we calculate the following distance metrics for any new sample for user i at time k, $\mathbf{x}_k(i)$, which is classified as normal by the global classifier:

$$R_i^{\max} = \max_{\mathbf{x}_j \in \mathcal{N}_i^k, \mathbf{x}_k \in \mathcal{N}_i^k} \{\sqrt{(\mathbf{x}_j - \mathbf{x}_k)^2}\}, \tag{3}$$

$$D_{\mathcal{X}}(\mathbf{x}_k(i)) = \operatorname*{median}_{\mathbf{x} \in \mathcal{X}} \{\sqrt{(\mathbf{x}_k(i) - \mathbf{x})^2}\}, \tag{4}$$

$$D_{\mathcal{N}}^{\max}(\mathbf{x}_k(i)) = \max_{\mathbf{x} \in \mathcal{N}_i^k} \{\sqrt{(\mathbf{x}_k(i) - \mathbf{x})^2}\}, \tag{5}$$

where $R_i^{\max}$, $D_{\mathcal{X}}(\mathbf{x}_k(i))$ and $D_{\mathcal{N}}^{\max}(\mathbf{x}_k(i))$, respectively, represents the maximum radius of the current aggregated normal cluster for user i, the median of distances from the current sample $\mathbf{x}_k(i)$ to samples in cluster $\mathcal{X}$, and the maximum distance of the current sample from the collected

normal samples. Then, we check the following conditions for each new sample $\mathbf{x}_k(i)$:

$$\begin{cases} D_{\mathcal{N}}^{\max}(\mathbf{x}_k(i)) \leq \alpha R_i^{\max}, \\ D_{\mathcal{N}}(\mathbf{x}_k(i)) < D_{\mathcal{X}}(\mathbf{x}_k(i)) \quad \text{for } \mathcal{X} \in \{\mathcal{S}, \mathcal{V}, \mathcal{F}\} \end{cases} \tag{6}$$

where α is a tuning parameter to make a balance between the classifier sensitivity and specificity. The first condition investigates the relative distance of the new sample from other normal samples, whereas the second condition verifies the closeness of the new sample to normal samples compared to its distance from abnormal samples. If these condition are satisfied, the sample is marked as *firm normal*. More formally, we set $f_2() : \mathbf{x}_k \mapsto \hat{y}_k^2 = N$, where $f_2()$ is the secondary classifier. If (6) does not hold, a yellow alarm is triggered as follows.

C. *Personalized Classifier*

For the samples which do not pass the conditions of (6), we perform a reclassification to identify the type of the yellow alarm. Note that each normal sample $\mathbf{x}_k(i)$, can be represented by a vector that originates from the current patient's normal cluster centroid $\mathbf{c}_N^k(i)$ to the sample, i.e.

$$\mathbf{v}_k(i) = \mathbf{x}_k(i) - \mathbf{c}_N^k(i) = \mathbf{x}_k(i) - \sum_{\mathbf{x} \in \mathcal{N}_i^k} \mathbf{x}/|\mathcal{N}_i^k|, \tag{7}$$

where $|\mathcal{N}_i^k|$ is the number of accumulated firm normal samples for user i up to time k. Likewise, we represent the centroid of 3 abnormal clusters $\mathcal{X} \in \{\mathcal{S}, \mathcal{V}, \mathcal{F}\}$ by $\mathbf{c}_{\mathcal{X}}$ and develop tendency vectors connecting the current sample $\mathbf{x}_k(i)$ to each of the abnormality classes as $\mathbf{v}_{\mathcal{X}}(i) = \mathbf{c}_{\mathcal{X}} - \mathbf{x}_k(i)$. The angles between the vectors $\mathbf{v}_k$ and $\mathbf{v}_{\mathcal{X}_j}$ can be used to determine the type of the yellow alarm as follows:

$$\hat{y}_k^2(i) = \underset{\mathcal{X} \in \{\mathcal{S}, \mathcal{V}, \mathcal{F}\}}{\operatorname{argmin}} \{d(\mathbf{v}_k(i), \mathbf{v}_{\mathcal{X}}(i))\} \tag{8}$$

for a given distance metric. For instance, we can use cosine distance to quantify the alignment as follows:

$$d(\mathbf{v}, \mathbf{w}) = 1 - \frac{\mathbf{v}^T\mathbf{w}}{|\mathbf{v}||\mathbf{w}|} = 1 - \frac{\mathbf{v}^T\mathbf{w}}{\sqrt{(\mathbf{v}^T\mathbf{v})(\mathbf{w}^T\mathbf{w})}} \tag{9}$$

The outputs of two global and local classifiers, $\hat{y}_k^1(i)$ and $\hat{y}_k^2(i)$, collectively determine the final label $y_k(i)$ as follows:

$$\hat{y}_k = \begin{cases} \text{Red alarm of type } \hat{y}_k^1 & \text{if } \hat{y}_k^1 \neq N \\ \text{Yellow alarm of type } \hat{y}_k^2 & \text{if } \hat{y}_k^1 = N \text{ and } \hat{y}_k^2 \neq N \\ N & \text{if } \hat{y}_k^1 = N \text{ and } \hat{y}_k^2 = N \end{cases} \tag{10}$$

Therefore, the final resulting label $y_k(i)$ belongs to $\mathcal{Y}_2 = \{N, S_r, V_r, F_r, S_y, V_y, F_y\}$, where subscripts r and y represent red and yellow alarms.

IV. Nonlinear Transformation

The essence of the deviation analysis is to assess the angles between the vectors connecting the centroid of normal cluster to abnormality clusters (i.e.$\mathbf{v}_{\mathcal{X}_j} = \mathbf{c}_{\mathcal{X}_j} - \mathbf{c}_{\mathcal{N}}$). Therefore, a symmetric geometry is a critical requirement. More specifically, we desire a maximal separation between these vectors. However, the geometry of clusters in the feature space Ω^d depends on the feature extraction stage $g()$. This geometry for the extracted features does not comply with this property as shown in Fig. 1. Here, we propose to use a nonlinear mapping to reshape the clusters to achieve the required symmetry, as illustrated in Fig. 2.

Our proposed approach is inspired by the nonlinear kernel method [17]. Here, we directly apply the nonlinear mapping to project each data vector $\mathbf{x}_k$ into a vector $\mathbf{z}_k$ which lies in a higher order dimensional space $\Phi^{d'} = \{\mathcal{X}_1', \mathcal{X}_2', \ldots, \mathcal{X}_C'\} = \{\mathcal{N}', \mathcal{S}', \mathcal{V}', \mathcal{F}'\}$, where $d' > d$. Here, we arbitrarily choose polynomial functions for the mapping purpose, as follows:

$$\begin{aligned} \mathbf{z}_k &= \mathbf{\Psi}_{\mathbf{w}}(\mathbf{x}_k) \\ &= ([w_1 \; w_2 \ldots w_{d'}] \circ [\psi_1(\mathbf{x}_k) \; \psi_2(\mathbf{x}_k) \ldots \psi_{d'}(\mathbf{x}_k)])^T \end{aligned} \tag{11}$$

where we fix the polynomial functions $\psi_i()$ and tune the coefficients w_i as optimization parameters. The goal of this optimization is to adjust mapping coefficients $\mathbf{w} = [w_1 \; \ldots \; w_d']^T$ to achieve the desired symmetric property for the projected data $\mathbf{z}_k$ while keeping the maximal separation among the clusters through objective functions:

$$o_1(\mathbf{w}) = \frac{1}{\min\limits_{c,d=2,\ldots,p \text{ and } c \neq d} \{d(\mathbf{v}_{\mathcal{X}_c}, \mathbf{v}_{\mathcal{X}_d})\}} \tag{12}$$

$$o_2(\mathbf{w}) = \frac{SW}{SB} = \frac{\sum_{c=1}^{C} \sum_{\mathbf{z} \in \mathcal{X}_c} (\mathbf{z} - \mathbf{c}_{\mathcal{X}_c})^T (\mathbf{z} - \mathbf{c}_{\mathcal{X}_c})}{\sum_{c=1}^{C} \sum_{d=1, d \neq c}^{C} (\mathbf{c}_{\mathcal{X}_c} - \mathbf{c}_{\mathcal{X}_d})^T (\mathbf{c}_{\mathcal{X}_c} - \mathbf{c}_{\mathcal{X}_d})}$$

By minimizing $o_1(\mathbf{w})$, we gain maximal symmetry for abnormal classes by maximizing the pairwise cosine distances among vectors $\mathbf{v}_{\mathcal{X}_c}$. On the other hand, $o_2(\mathbf{w})$ quantifies the ratio of within cluster variance to between-cluster variance, hence controls the separation of the clusters. The goal of the optimization problem here is to jointly minimize $o_1(\mathbf{w})$ and $o_2(\mathbf{w})$ subject to $|\mathbf{w}|_2 = 1$. This is a multi-objective non-convex optimization problem. Here, we use multi-objective particle swarm optimization (MOPSO) for its fast convergence and superior performance to other methods [9]. MOPSO approximates the Pareto front of the objective functions. With two objective functions, the Pareto optimal particles can be plotted as a curve in 2-D space. In order to compare the improvement of objective functions from the original space Ω^d to the designed nonlinear space $\Phi^{d'}$, the algorithm is tested on a 2-D simulated data, as shown in Fig.1. The details of the developed MOPSO is omitted here for the sake of brevity.

Fig.3 shows that the estimated Pareto font when using the transformed data $\mathbf{z}_k$ significantly dominates the Pareto front using the original data $\mathbf{x}_k$. The kernel used in Pareto front comparison is a 3-degree polynomial kernel. The result shows that using nonlinear transformation with optimized

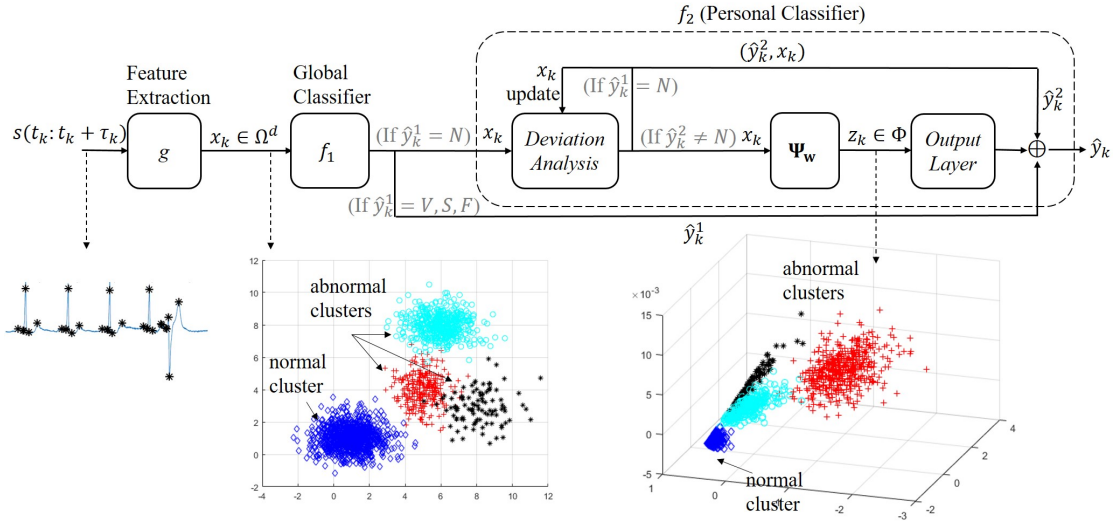

Fig. 1: Conceptual flow chart for nonlinear reshaping system

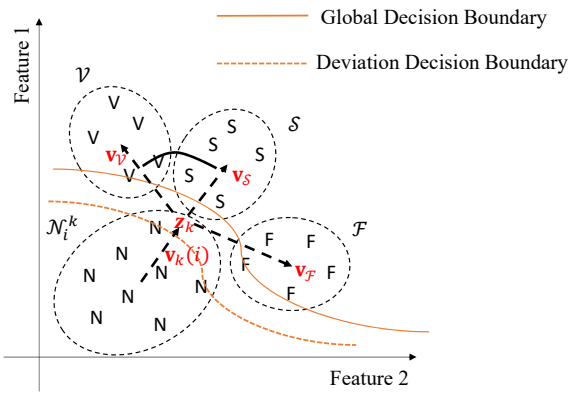

Fig. 2: A conceptual illustration for data spatial arrangement after nonlinear reshaping

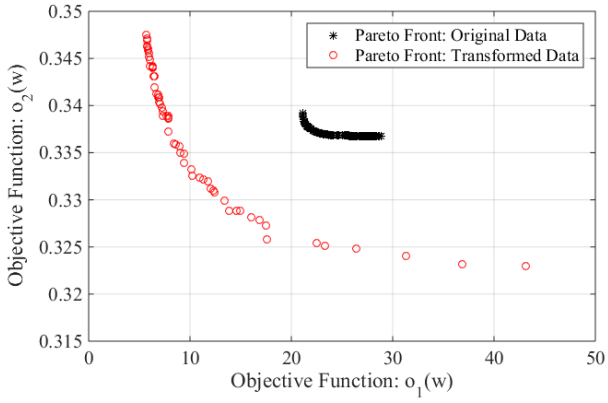

Fig. 3: Estimated Pareto front for before and after nonlinear transformation.

coefficients provides sufficient flexibility of improving different optimization metrics. The same conclusion can be drawn from Fig.1, which illustrates the cluster geometry before and after the proposed nonlinear transformation for a polynomial kernel of power 2 for a simulated data.

V. PREDICTIVE MODELING

An important utility of yellow alarms is their ability to predict upcoming abnormalities, as we conjecture that the probability of observing a red alarm of a specific type increases after observing a yellow alarm of the same type. More formally, if $\hat{y}_k = X_y$ is a yellow alarm of type X, and $\hat{y}_{k+i}, i > 0$ is the first subsequent red alarm, we have

$$P(\hat{y}_{k+i} = X_r | \hat{y}_k = X_y) > P(\hat{y}_{k+i} = X_r) \qquad (13)$$

This conjecture is confirmed by numerical results presented in section VI.

VI. RESULTS

In this section, numerical results using MIT-BIH arrhythmia database (MITDB) [10] are provided to show the performance of the proposed method. Here, we firstly map the original 22-dimensional feature vectors representative of cardiac segments into a 8-dimensional vectors $\mathbf{x}_{8\times1}$ using principal component analysis (PCA). In order to realize the nonlinear transformation in (11), we use polynomial function of order 3. Taking the computational cost into account and also to avoid over-fitting of higher-order data samples, we randomly discard some cross terms $(x_i x_j^2, i \neq j)$ and keep only 32 terms in $\mathbf{z}_{32\times1} = \Phi_{\mathbf{w}}(\mathbf{x}_{8\times1})$. Table III, shows the performance of the proposed method in classification of ECG signal segments. In order to average the results over all recordings, we present the median, interquartile range (IQR), mean and standard deviation of accuracy (AC), sensitivity (SE) and specificity (SP). The results are promising and the median of accuracy for all classes are in the range of $88\% - 99\%$. Sensitivity and specificity of the proposed method exhibits similar ranges. The mean accuracy is at least

86% excluding class V. Therefore, this system is very less likely to miss an important alarm or to report false alarms.

TABLE III: Classification results of the proposed method.

Class N	median(%)	IQR(%)	mean(%)	std (%)
AC	94.8	19.52	86.62	18.55
SE	97.21	17.36	87.47	19.26
class V	median(%)	IQR(%)	mean(%)	std (%)
AC	86.11	27.54	76.41	22.81
SP	99.71	11.22	90.18	18.52
class S	median(%)	IQR(%)	mean(%)	std (%)
AC	99.28	2.24	98.29	2.57
SP	99.64	22.17	97.56	6.06
class F	median(%)	IQR(%)	mean(%)	std (%)
AC	97.91	8.2	93.85	7.84
SP	100.00	0.03	99.12	3.6

More importantly, the predicting ability of the proposed method is worthy of evaluating separately. In order to quantify the conditional probability of observing a red alarm after a preceding yellow alarm of similar type in (13), we count the number of predicted samples as follows:

$$P(\hat{y}_{k+i} = X_r | \hat{y}_k = X_y) = \frac{\text{\# of } y_{k+i} = X \text{ after } \hat{y}_k = X_y}{\text{\# of true alarms after } \hat{y}_k = X_y}$$
$$P(\hat{y}_{k+i} = X_r) = \frac{\text{\# of true alarm of type } X \ (y_k = X)}{\text{\# of all true alarms}} \tag{14}$$

The summary of results for all 21 test records is presented in Table. IV. The last 4 columns of the Table. IV show the probability of having a subsequent true abnormality of any type after observing a yellow alarm. These results confirm the predictive power of yellow alarms. For instance, the absolute probability of observing a segment with abnormal classes V, S, and F is respectively $\frac{96}{96+29+18} = 67\%$, $\frac{29}{96+29+18} = 20\%$ and $\frac{18}{96+29+18} = 13\%$, based on their relative frequencies. However, these probabailities after observing a yellow alarm of type Vp are respectively $\frac{38}{38+11+2} = 75\%$, $\frac{11}{38+11+2} = 21\%$ and $\frac{2}{38+11+2} = 4\%$. This means that the probability of observing a red alarm of type V is $75\% - 67\% = 8\%$ higher than its absolute probability. The same trend holds for other yellow alarms as well. The results suggest a more in-depth study of the concept of yellow alarms for heart monitoring.

VII. CONCLUSIONS

In this paper, we proposed a novel patient-specific ECG signal analysis for remote heart monitoring. A key novel property of this method is its multi-resolution alarm reporting capability. For instance, yellow alarms can be triggered for normal cardiac circles which show considerable deviations from the patient-specific normal morphology. The deviation analysis is realized by a novel nonlinear transformation with controlled coefficients using MOPSO algorithm. The results verify the accuracy of the proposed method with a classification accuracy range of $88\% - 99\%$ for different ECG records in publicly available MIT-BIH database. A key unique feature of the proposed method is the utility of yellow alarms in predicting upcoming red alarms, which can help the patients to take required precautions (such as taking rest, taking prescribed medications, and contacting their heath providers). The prediction improvement is in the range of 10%. The proposed methodology is general and can be used for other biomedical signals as well.

TABLE IV: Predictive power of yellow alarms

	# of red alarms after yellow alarms				Probability of next abnormality (%)			
True Alarm	V_y	S_y	F_y	Total	V_y	S_y	F_y	Total
True V	38	23	35	96	75	75	61	67
True S	11	10	8	29	21	29	14	20
True F	2	2	14	18	4	6	25	13

REFERENCES

[1] S. H. Jambukia, V. K. Dabhi, and H. B. Prajapati, "Classification of ECG signals using machine learning techniques: A survey," in *ICACEA 2015*. IEEE, 2015, pp. 714–721.

[2] A. Dallali, A. Kachouri, and M. Samet, "Fuzzy c-means clustering, neural network, wt, and hrv for classification of cardiac arrhythmia," *ARPN Journal of Engineering and Applied Sciences*, vol. 6, no. 10, p. 2011, 2011.

[3] Z. Zidelmal, A. Amirou, D. Ould-Abdeslam, and J. Merckle, "ECG beat classification using a cost sensitive classifier," *Computer methods and programs in biomedicine*, vol. 111, no. 3, pp. 570–577, 2013.

[4] T. Ince, S. Kiranyaz, and M. Gabbouj, "A generic and robust system for automated patient-specific classification of ECG signals," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1415–1426, 2009.

[5] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 9, pp. 891–900, Sept 1997.

[6] W. Jiang and S. G. Kong, "Block-based neural networks for personalized ECG signal classification," *IEEE Transactions on Neural Networks*, vol. 18, no. 6, pp. 1750–1761, 2007.

[7] F. Afghah, A. Razi, S. R. Soroushmehr, S. Molaei, H. Ghanbari, and K. Najarian, "A game theoretic predictive modeling approach to reduction of false alarm," in *International Conference on Smart Health*. Springer, 2015, pp. 118–130.

[8] J. Chen, H. Peng, and A. Razi, "Remote ECG monitoring kit to predict patient-specific heart abnormalities," *Journal of Systemics, Cybernetics and Informatics*, vol. 15, no. 4, pp. 82–89, 2017.

[9] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Transactions on evolutionary computation*, vol. 8, no. 3, pp. 256–279, 2004.

[10] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[11] P. de 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, July 2004.

[12] B. N. Singh and A. K. Tiwari, "Optimal selection of wavelet basis function applied to ECG signal denoising," *Digital signal processing*, vol. 16, no. 3, pp. 275–287, 2006.

[13] Z. Zidelmal, A. Amirou, M. Adnane, and A. Belouchrani, "QRS detection based on wavelet coefficients," *Computer methods and programs in biomedicine*, vol. 107, no. 3, pp. 490–496, 2012.

[14] G. B. Moody, "Evaluating ECG analyzers," *WFDB Applications Guide*, 2003.

[15] T. Mar, S. Zaunseder, J. P. Martínez, M. Llamedo, and R. Poll, "Optimization of ECG classification by means of feature selection," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 8, pp. 2168–2177, Aug 2011.

[16] A.-A. EC57, "Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms," *Association for the Advancement of Medical Instrumentation, Arlington, VA*, 1998.

[17] L. Wang, A. Razi, M. Rodrigues, R. Calderbank, and L. Carin, "Nonlinear information-theoretic compressive measurement design," in *International Conference on Machine Learning*, 2014, pp. 1161–1169.