

Design of a Low-Power On-Body ECG Classifier for Remote Cardiovascular Monitoring Systems

Taihai Chen, Evangelos B. Mazomenos, Koushik Maharatna *Member, IEEE*, Srinandan Dasmahapatra and Mahesan Niranjan

Abstract—In this paper we first present a detailed study on the trade-off between the computational complexity (directly related to the power consumption) and classification accuracy for a number of classifiers for classifying normal and abnormal ECGs. In our analysis we consider the spectral energy of the constituent waves of the ECG as the discriminative feature. Starting with the exhaustive exploration of single heart-beat based classification to ascertain the complexity-accuracy trade-off in different classification algorithms, we then extend our study for multiple heart-beat based classification. We use data available in Physionet as well as samples from Southampton General Hospital Cardiology Department for our investigation. Our primary conclusion is that a classifier based on Linear Discriminant Analysis (LDA) achieves comparable level of accuracy to the best performing Support Vector Machine (SVM) classifiers with advantage of significantly reduced computational complexity. Subsequently, we propose an ultra low-power circuit implementation of the LDA classifier that could be integrated with the ECG sensor node enabling on-body normal and abnormal ECG classification. The simulated circuit is synthesized at 130 nm technology and occupies 0.70 mm² of silicon area (0.979 mm² after Place and Route) while it consumes 182.94nW @ 1.08 V, estimated with Synopsys PrimeTime when operating at 1 KHz. These results clearly demonstrate the potential for low-power implementation of the proposed design.

Index Terms—ECG Classification, Low-Energy, Discrete Wavelet Transform, Computational Complexity Analysis, Remote Healthcare Applications

I. INTRODUCTION

RECENT advances in the Internet of Things paradigm (IoT) have paved the way for developing new services by the seamless interconnection of a number of heterogeneous devices and the internet in both wired and wireless fashion making information available “anytime anywhere”. One such high potential service is the development and mass deployment of next-generation remote healthcare systems that enable continuous monitoring of chronic disease patients and therefore facilitating clinicians in initiating preventive intervention even before the symptoms of a critical episode are

fully manifested. Such “proactive” rather than the traditional “reactive” approach may not only reduce the mortality rate significantly, but also result in significant cost savings by minimising the hospital admission rate, bed time and costly human intervention. The fundamental principle behind it is to combine vital sign data of the patients captured by appropriate sensors on body with patient’s history and practical clinical knowledge for clinical decision, based on which an alarm is generated for a possible impending episode. Collected data can then be transmitted to an proper facility for further analysis.

It is already well-known in the clinical community that continuous variability analysis of the vital sign data provides more enriched clinically important information than a discrete “snap shot” of them. Hence, continuous monitoring and the associated signal processing in nomadic environment are the key aspects of the next-generation remote healthcare systems. The traditional approach is to continuously transmit the vital sign data, captured by wireless sensors, to a server where the computationally intensive signal analysis task takes place. This is done due to the fact that the resource constrained nature of the sensors does not allow the execution of complex signal processing routines. However, the energy expenditure of the front-end radio system does not permit continuous monitoring for long time [1]. In addition, since the system’s main purpose is to generate an alarm through preliminary analysis, a trade-off between computational complexity and accuracy of the signal processing algorithms can be made to achieve a low-power on-sensor implementation. Thus, unless any abnormality is detected data transmission is not required. This in turn negates the requirement for continuous use of the front-end radio system. The analysed data can be stored in the local memory of the sensor and be transmitted at a pre-set interval in burst mode, therefore maintaining the notion of continuous monitoring while saving significant energy [1].

One of the main application areas of the IoT concept is on next-generation remote healthcare monitoring systems with a focus on Cardiovascular Diseases (CVD), which according to the World Health Organisation (WHO) is the most prominent disease [2]. The Electrocardiogram (ECG) captures the electrical activity of the heart as a time series data and acts as the first screening tool to detect heart abnormalities in the standard clinical practice. Subsequently, the ECG plays the central role in developing a remote CVD monitoring system as it is portable and widely available. In such a system, the main role of the ECG is to classify the normal and abnormal heart rhythm and accordingly produce an alarm. Any attempt for specific disease diagnosis could be hazardous,

Manuscript received August 15, 2012; revised Oct 29, 2012; accepted Dec 23, 2012. This work was supported by E.U. ARTEMIS Joint Undertaking under the Cyclic and person-centric Health management: Integrated appRoach for hOme, mobile and clinical eNvironments - (CHIRON) Project, Grant Agreement # 2009-1-100228.

T. Chen, E. B. Mazomenos, K. Maharatna, S. Dasmahapatra and M. Niranjan are with School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK. (e-mail: tc10g09, ebm, km3, sd, mn@ecs.soton.ac.uk)

Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

due to the limited amount of information (small number of leads) available on a remote CVD monitoring system and the various co-morbidities or confounding conditions that manifest themselves in a similar way in the ECG trace albeit being different clinical conditions. For instance, some ECG morphological changes, indicative of some diseases, may not be captured as they can only be seen in particular leads, which may not be available in the remote system. Moreover, various heart diseases are known to exhibit similar morphological changes of the QRS complex (e.g. QRS duration) [3], thus simply considering these as features and trying to classify a disease based on them, is not practical as this in principle will result into a one-to-many mapping. In addition, even in clinical settings, where standard 12/15 lead ECG recordings are available, the ECG is only used as the first screening tool to help clinicians decide what elaborate clinical investigation (medical imaging, stress test etc..) is further required. Subsequently after consultation with cardiologists from Southampton General Hospital [4] and Policlinico Umberto I [5] participating in the E.U. CHIRON project [6], the clinical expectation for a remote CVD system is to indicate if there is any heart abnormality, irrespective of the specific condition causing the abnormality.

In this paper we first carry out a thorough exploration on the trade-off between the computational complexity and accuracy of a number of classifiers in classifying normal and abnormal ECGs with an aim to ascertain which is optimal for on-sensor classification. This is considered to be the major novelty of this work, as no comprehensive study on such trade-off has been presented with particular focus on mobile monitoring systems. The initial exploration is performed considering a single-beat-scenario (one simultaneously captured beat per participating lead) from each ECG record and is then extended to multiple beats per participating lead (multiple-beat-scenario). Our investigation shows that Linear Discriminant Analysis (LDA) exhibits the best trade-off. Accordingly, we present a novel circuit implementation of the LDA classifier which is suitable for on-body sensor classification of normal and abnormal ECGs with low power. The proposed architecture is not yet implemented as a standalone ASIC, since it is a subsystem of a complete on-body wearable CVD monitoring platform, currently under integration. This includes various additional subsystems (ECG data acquisition, ECG feature extraction, etc.) and is developed under CHIRON project [6], of which this work is a part. The rest of the paper is structured as follows: in Section II we provide the fundamental principles of the classifiers along with their computational complexity analysis and Section III describes the feature space generation. Results from classification experiments on both single-beat-scenario and multiple-beat-scenario are described in Sections IV and V respectively. Based on these outcomes, a low-power architecture for LDA is described in Section VI and its implementation and verification results are discussed in Section VII. Conclusions are drawn in Section VIII.

II. FUNDAMENTALS OF THE CLASSIFIERS AND THEIR COMPLEXITY

Although many different approaches have been considered in ECG classification [7]–[13], in this work we consider five

different classifiers which have been shown in several studies to have high performance. These are, Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA) [14], Support Vector Machine (SVM) with Linear (SVM_L) and Quadratic (SVM_Q) kernels [7] and k-Nearest Neighbours (k-NN) [8]. In the following we provide brief principles of these classification techniques and calculate their arithmetic computational complexity.

The fundamental assumption in LDA is that every class distribution is Gaussian in nature and the covariance matrices of the classes are identical. On the other hand QDA assumes that the classes follow Gaussian distributions but with different covariance matrix for each class [15]. The associated decision boundary of LDA and QDA is given as $f(\mathbf{x}) = \sum_{i=1}^n L_i x_i + b$ and $f(\mathbf{x}) = \sum_{i,j=1}^n Q_{ij} x_i x_j + \sum_{i=1}^n L_i x_i + b$ respectively, where Q_{ij} , L_i and b indicate the coefficients and the interception of the hyperplanes respectively.

SVMs belong to the class of the binary classifiers based on maximum margin strategy [16]. For determining the separating hyperplane in high-dimensional feature space SVM uses a kernel $K(\cdot)$. Typically a linear or a quadratic kernel is used for the SVM-based classification. The output of SVM is computed from $y(\mathbf{x}) = \text{sgn}(\sum_k \alpha_k y(\mathbf{x}_k) K(\mathbf{x}_k, \mathbf{x}) + b)$, where \mathbf{x} , sgn , k and α_k represent the new sample for labelling, sign function, the number of support vectors (x_k) and the Lagrangian multipliers respectively. For Linear (SVM_L) and Quadratic kernel (SVM_Q), the associated kernel functions [16] are $K_L = \langle \mathbf{x}_k, \mathbf{x} \rangle$ and $K_Q = (\langle \mathbf{x}_k, \mathbf{x} \rangle + 1)^2$, where $\langle \cdot \rangle$ denotes inner product operation.

k-NN is a nonparametric classifier where every new sample is labeled from the majority class which has k-nearest neighbours around this sample in the training set. The distance between the new sample and the training samples is calculated using the Euclidean norm given as $d(\mathbf{x}, \mathbf{x}_k) = \sqrt{\sum_{i=1}^n (x_i - x_{k_i})^2}$, where \mathbf{x} , \mathbf{x}_k indicate the new sample and the training data respectively. In this work, we set $k = 3$.

A. Comparison of arithmetic computational complexity

Typically there are two sets of computations associated with every classification techniques - computations required during training and computations required for labelling new sample. In practice, the first part is a one-time offline procedure that is carried out before the deployment of the classifier. On the other hand, sample labelling is the actual computational procedure taking place during the operation of the classifier. Therefore in the computational complexity analysis we have considered only this part. The computational complexity for each method is expressed in terms of their required arithmetic operations as this is representative of the energy consumption required for each classifier. Since several implementations of the same arithmetic function and the classifier architecture are possible, to provide a uniform platform, we consider flat unfolded architecture without any resource sharing or parallelism and describe the numbers of arithmetic operations required for each of them. For k-NN, apart from the arithmetic operations involved in computing the distance, those for determining the k-nearest neighbours from the full set of training samples are

also considered. For simplicity, the complexity of a subtraction operation is considered equal to that of an addition. The arithmetic complexities of different classifiers are illustrated in Table I where N , M , S indicate the dimension of feature vector, the number of Support Vectors (SVs) in SVM and the number of training samples respectively.

TABLE I
TOTAL NUMBER OF ARITHMETIC OPERATIONS INVOLVED IN LABELLING A
NEW SAMPLE FOR THE FIVE CLASSIFIERS

	Add (+)	Mul (\times)	Sqr ($()^2$)	Sqrt ($\sqrt{}$)
LDA	N	N	0	0
QDA	$N^2 + N$	$N^2 + N$	N	0
SVM_L	$(N+1)M - 1$	$(N+2)M$	0	0
SVM_Q	$(N+2)M - 1$	$(N+2)M$	M	0
k-NN	$2S(N+1) - 6$	0	SN	S

From Table I it is evident that for a given dimension of the feature vector LDA exhibits the least computational complexity whereas SVM and k-NN strongly depends on the number of SVs and training samples used respectively. In addition, an increase in the number of feature vectors will cause a linear increase in the computational complexity of LDA, while the rest exhibit an approximately quadratic increase (considering M and S comparable to N and squaring operation equivalent to multiplication). It is to be noted that as an arithmetic block a multiplier (and a divider) is several times more energy consuming than an adder. Therefore due to smaller number of multipliers required in LDA, it is expected that LDA will consume much less energy than the other classifiers.

In order to create an unified metric describing the overall computational complexity for each of the classifiers, we used 2-input NAND gate complexity (NG). Considering unfolded architecture and no resource sharing for each of the arithmetic computational modules, and b -bit wordlength implementation, the number of transistors required for each operation is $T_+ = 24b$, T_\times and $T_{()^2} = 30b^2 - 36b$, and $T_{\sqrt{}} = 18(\frac{b}{2} + 1)(\frac{b}{2} + 3)$ [17], where T_* denotes the transistor count for the arithmetic operation (*). Since a 2-input NAND gate requires 4 transistors, these numbers could be transformed into NAND gate equivalent as $G_+ = 6b$, G_\times and $G_{()^2} = \frac{15}{2}b^2 - 9b$ and $G_{\sqrt{}} = \frac{9}{2}(\frac{b}{2} + 1)(\frac{b}{2} + 3)$.

III. FEATURE SPACE GENERATION

In standard clinical settings, 12-lead ECG is used where each lead provides a different ECG trace. However, in a remote CVD systems, the deployment of 12-lead ECG is not feasible. In general, at maximum 5 leads are used in such systems [18]. We therefore restrict ourselves in exploring the performance of the classifiers considering no more than 5 leads.

In this work we have opted to explore the use of the spectral energy of the ECG signal as the main feature for classification. However instead of considering the spectral energy of the entire ECG-beat, we use the spectral energy of specific parts of the ECG-beat as separate features. This is done from the point of view that different classes of ECG abnormalities may be reflected in different ECG waves (P, QRS and T) therefore the individual characteristics of each of them may

have more discriminative properties for classifying normal and abnormal ECG compared to the spectral energy of the entire PQRST complex. We have performed an extensive exploration on the applicability of the spectral energy as a discriminative feature and it turned out that it can indeed be considered as a robust feature particularly against misdetections of the ECG boundaries - a typical situation often encountered in automated ECG analysis. However the detailed description of that exploration is beyond the scope of this work and will be described in our future communication.

To calculate the spectral energy of each constituent ECG wave, we first consider an isolated PQRST complex where the ECG segmentation algorithm described in [19] is applied to extract the temporal boundaries of each wave. Following the PQRST complex is subjected to Discrete Wavelet Transform (DWT) with Haar as the basis function. Without any loss of generality other forms of basis functions could also be used. The time-frequency localisation property of DWT is utilised here for isolating the DWT coefficients corresponding to each of the P, QRS and T wave. It has already been shown that the high-frequency components of the ECG signal, like the QRS complex are better localised at the DWT decomposition levels 2 and 3 [20] whereas the low-frequency components like P and T waves are better localised at level 5. The DWT is implemented using Mallat's algorithm [21] where at each level of decomposition two coefficient vectors - one detail (cD) and one approximate (cA) - are generated. The cA at i^{th} level is used as the input of $(i^{th} + 1)$ decomposition level. In our exploration we exclude the DWT coefficients at level 1 as from our experiments it has been shown that this level is mostly dominated by high-frequency noise and therefore contains little information pertinent to the ECG signal itself. Once DWT coefficients are generated, the spectral energy for a given interval $[n_1, n_2]$ is calculated using Equation (1) below, where $W_{m,n}$ denotes the cD vector at decomposition level m .

$$E_m^{(n_1, n_2)} = \sum_{n=n_1}^{n_2} |W_{m,n}|^2 \quad (1)$$

We have opted to utilize the spectral energy of the following ECG parts as our feature set, the QRS complex, the P, T wave, the QT interval and the PR interval. The QRS spectral energy is obtained from both level 2 and 3 coefficients leading to two different features QRS_2 and QRS_3 . The energy of the P, T wave and PR interval is calculated from level 5 coefficients resulting in three more spectral energy features - P_5 , T_5 , PR_5 . Finally the QT interval spectral energy, which contains both high and low frequency components, is calculated by summing the individual QT interval spectral energies from coefficients in level 3 and 5, as well as coefficients in level 3, 4 and 5 to produce two different calculations of the QT interval spectral energy (QT_{35} , QT_{345}). The reason for including level 4 in QT_{345} is to capture the transition of the spectral energy from high to low frequency. In total 7 distinct spectral energy features are calculated on a per lead basis and grouped into three categories - the low-frequency feature group (L): P_5 , T_5 , PR_5 ; the high-frequency feature group (H): QRS_2 , QRS_3 ; and the combined high and low-frequency feature group (B):

QT₃₅, QT₃₄₅.

IV. CLASSIFIER COMPLEXITY-ACCURACY TRADE-OFF FOR SINGLE-BEAT-SCENARIO CLASSIFICATION

In this section, we present an analysis on the complexity-accuracy trade-off of the five methods considered for the classification of normal and abnormal ECGs based on the single-beat-scenario. The spectral energy features are extracted on a per lead basis, which means that if N leads are available N simultaneously captured heart-beats are used.

A. ECG Databases

Two ECG databases were used to provide 104 12-lead records as our experimental set with full clinical diagnosis/annotations. Within these 104 records, 52 are categorised as normal and 52 as abnormal. Specifically, all of the 52 normal records (healthy control) are selected from the PTB database (PTBDB), available in the PhysioNet [22], which consists of 549 standard 15-lead ECG records, from healthy and various disease categories subjects, sampled at 1 KHz. For the set of abnormal samples, 17 abnormal records are obtained from PTBDB, which cover all available Myocardial Infarction subclasses (anterior, inferior, lateral and posterior), 18 are equally collected from the other six disease classes (e.g. cardiomyopathy, bundle branch block, dysrhythmia, myocardial hypertrophy, valvular heart disease and myocarditis) and the rest 17 abnormal records are obtained from the Southampton General Hospital Cardiology Department's database [4] and pertain to patients with myocardial scar. These records are standard 12-lead ECG sampled at 500 Hz, of 10 seconds length.

B. Feature Ranking and Selection

We begin our exploration using all 12-lead ECGs available and extract one full PQRST complex from each of the leads. Each of these isolated PQRST complex undergoes the feature generation procedure described in Section III resulting in 7 distinct spectral energy features per ECG beat and totalling to $7 \times 12 = 84$ features. Our aim is to identify the best set of features, through simulation using different lead combinations, and evaluate the classification performance under the five aforementioned classifiers. Since exhaustive simulation in this large feature space is time demanding, Fisher's criterion [23] is employed to select one feature from the L, H and B frequency groups (described in Section III) for each lead, which can separate the two classes to the maximal extent. By doing so, we expect to identify the most discriminating features for classification. In essence, Fisher's criterion calculates the ratio of the between-class variance to the within-class variance on the basis of one feature and indicates the extent of mean separation and overlap between the two classes. Once the ratios of the 84 features are obtained, the distinctive feature for each feature frequency group is determined by selecting the highest one within each frequency group of each lead. The final selected features for each lead, after applying this principle, are shown in Table II.

TABLE II
THE MOST DISCRIMINANT ENERGY FEATURE IN EACH FREQUENCY GROUP FOR EACH LEAD

	Lead (Abbr)	I (1)	II (2)	III (3)	aVR (4)	aVL (5)	aVF (6)
Fea Grp	L	T ₅	T ₅	T ₅	T ₅	T ₅	T ₅
	H	QRS ₂	QRS ₃	QRS ₂	QRS ₃	QRS ₃	QRS ₂
	B	QT ₃₄₅	QT ₃₄₅	QT ₃₅	QT ₃₄₅	QT ₃₅	QT ₃₄₅
	Lead (Abbr)	V1 (7)	V2 (8)	V3 (9)	V4 (10)	V5 (11)	V6 (12)
Fea Grp	L	P ₅	P ₅	PR ₅	T ₅	T ₅	T ₅
	H	QRS ₃	QRS ₂	QRS ₂	QRS ₃	QRS ₃	QRS ₂
	B	QT ₃₄₅	QT ₃₅	QT ₃₅	QT ₃₅	QT ₃₅	QT ₃₄₅

After deriving the optimal features, we use exhaustive simulation to identify the best combination of leads, that results in maximum classification accuracy. Considering l number of leads being available, out of the total of 12 leads, the number of possible lead combination is C_l^{12} , where C_q^p denotes the combination of order q from p elements. In each of these lead combinations, we opt to select at least one (at most three) feature(s) from each individual lead and build the feature space. We restrict ourselves in considering at maximum 5 leads, in order to remain in line with the constraints imposed by the application scenario of remote CVD monitoring. The decision to classify the ECG record to a class is based on the features from the l available leads (heart-beats). All samples from each feature are normalized with respect to their mean and standard deviation at the beginning of our exploration. The training data has been used for optimizing the parameters in the parametric models of the classifiers while the regularization parameter C_s of SVM is set to 1. Conventional quadratic programming solving method is selected in the training phase for SVM.

For evaluating the performance of each of the classifiers we use the metrics of specificity, sensitivity and overall testing accuracy as defined in Eq. 2, where TP and FP denote True Positives (correct abnormal classifications) and False Positives (wrong abnormal classifications), while TN and FN denote True Negatives (correct normal classifications) and False Negatives (wrong normal classifications). Finally, NR refers to the total number of records (104 in our experiments).

$$\begin{aligned}
 Spe &= \frac{TN}{(TN + FP)} \times 100\% \\
 Sen &= \frac{TP}{(TP + FN)} \times 100\% \\
 Acc &= \frac{(TP + TN)}{NR} \times 100\%
 \end{aligned} \tag{2}$$

By running exhaustive simulation and evaluating the results for each of the classifiers under consideration, we select the optimal lead combination and its associated feature combinations exhibiting maximum accuracy for every lead scenario. The term lead scenario refers to the number of participating leads. Initially only one lead $l = 1$ is considered and from the 12 available, the one that achieves maximum accuracy is selected for each classifier. Following, we keep this lead and combine it with each one of the remaining 11 in order to identify the best lead combination for $l = 2$. This process is repeated up to $l = 5$. By doing so, we ensure that the

performance of a classifier is improving as we add extra leads, hence their associated distinctive features. Table III shows the lead combinations with which the classifiers obtain maximal performance in all lead scenarios under consideration. The '(Lead, Feature) Combination' column shows which features from Table II in each feature frequency group for each lead are used.

TABLE III
FEATURE SPACE SELECTION

	# of Lead	(Lead, Feature) Combination
LDA	1	(3,LHB)
	2	(2,LH),(3,LB)
	3	(2,LHB),(3,LHB),(7,LB)
	4	(2,LHB),(3,LHB),(7,LB),(8,L)
	5	(1,L),(2,LB),(3,LHB),(7,B),(8,L)
QDA	1	(4,LB)
	2	(3,LHB),(4,LHB)
	3	(3,LH),(4,LB),(5,LHB)
	4	(3,LHB),(4,LB),(5,LHB),(10,H)
	5	(3,LH),(4,HB),(5,HB),(6,L),(10,LH)
SVM_L	1	(2,LH)
	2	(2,L),(3,LHB)
	3	(2,L),(3,LH),(8,LHB)
	4	(2,L),(3,LHB),(5,HB),(8,LB)
	5	(2,L),(3,LB),(5,B),(6,LHB),(8,LHB)
SVM_Q	1	(4,LH)
	2	(4,L),(8,LHB)
	3	(4,L),(5,LHB),(8,LHB)
	4	(3,LH),(4,L),(5,L),(8,LHB)
	5	(2,L),(3,LB),(4,L),(5,L),(8,LHB)
k-NN	1	(4,L)
	2	(4,L),(5,L)
	3	(4,LHB),(5,LHB),(9,H)
	4	(4,LHB),(5,LH),(8,B),(9,HB)
	5	(3,LHB),(4,LB),(5,LHB),(8,B),(9,H)

C. Classification Results Analysis and Discussion

To calculate the complexity of each classifier in terms of 2-input NAND gate equivalent (NG), a word length of $b = 16$ bit is considered. 10 runs of 10-fold cross validation is used to obtain consistent performance of the classifier. Table IV lists the the classification performance of LDA, QDA and k-NN for their corresponding best lead combination in different lead scenarios, along with their associated computational complexity. By adding features from extra leads, the overall testing accuracy of these three classifiers increases gradually. At the same time NG also grows. This is because the dimension of the feature vector increases.

Unlike the above three classifiers, the computational complexity and performance of SVMs are affected by the number of SVs employed. To increase/decrease the number of SVs, the regularisation parameter C_s can be tuned [7]. This indirectly influences both the performance and computational complexity of SVM classifiers. Thus, to investigate the way the performance and complexity of SVMs are affected by C_s , two cases are considered. Case 1 is when the maximum accuracy is obtained without being concerned about the number of SVs deployed, while Case 2 is the scenario where the minimum number of SVs, thus minimum complexity and the associated accuracy are obtained. Initially, C_s was set to 1, to find out the best lead (and feature) combination. Now, C_s is set after performing grid search ($C_s = 2^i$, $i = -15, -14 \dots 14, 15$)

during the training phase and the C_s^{min} value that achieves the minimum number of SVs is preserved for every lead scenario as this corresponds to the case of least computational complexity as shown in Table I.

Table V shows the performance and complexity analysis for Case 1. As we expect, both SVM_L and SVM_Q demonstrate increasing testing accuracy with respect to the number of leads deployed. Additionally with increasing number of leads the number of SVs decreases. This is expected as adding more leads increases the size of the feature set, which also leads to the maximum accuracy to be obtained with less SVs. Table VI depicts Case 2 for SVM. The obtained testing accuracy for SVM_L and SVM_Q shows that these two classifiers still achieve a similar level of performance to LDA and k-NN while outperforming QDA as in Case 1. The minimum number of SVs, as obtained with use of C_s^{min} is listed in Table VI. The difference in the number of SVs, between Case 1 and Case 2 for SVM_L is smaller than SVM_Q, where a reduction up to 30% can be observed when more than 1 leads are considered. This dramatic decrease in the number of SVs results in a reduction in the computational complexity of the SVM_Q compared to Case 1. As a result, in Case 2 the total complexity in terms of NG is reduced significantly only for SVM_Q compared to the NG in Case 1. Therefore, we conclude that the optimisation of the C_s parameter has a meaningfully positive effect only on the complexity of SVM_Q. When we compare the performance of the five classifiers for different lead scenarios, it is observed that SVM_Q (Case 1) exhibits the best performance in all situations except k-NN performs slightly better in 5 lead scenario. In addition, specificity tends to be higher than sensitivity in LDA, k-NN and SVM_Q, whereas in QDA the opposite happens. For SVM_L, the two metrics outperform each other depending on the lead scenario. Also, LDA outperforms QDA in every occasion and achieves comparable performance to SVM_Q.

By observing the overall NG required from Table IV, Table V and Table VI, for each of the classifiers to label a new sample, we argue that LDA approximately requires two orders of magnitude less NG than k-NN and SVM and one less than QDA. As the number of lead increases, NG increases gradually for LDA, QDA and k-NN while it remains fairly constant for both SVM_L and SVM_Q. The results for both accuracy and complexity indicate that while the NG count of LDA is significantly smaller, than that of the other classifiers, the accuracy it achieves is either the highest or within a 4% margin of the best classifier for all lead scenarios apart from the 1 lead scenario where LDA demonstrates a 7% less accuracy than SVM_Q.

V. CLASSIFIER PERFORMANCE FOR MULTIPLE-BEAT-SCENARIO CLASSIFICATION

Having performed the trade-off analysis for the single-beat-scenario, we extend our study considering multiple heart-beats to be utilized from each available lead. The primary reason for this exploration is that from a morphological point of view, relying only on a single heart-beat per lead to decide the class of the record, may result in misclassifications as it

TABLE IV
SPECIFICITY, SENSITIVITY, ACCURACY AND ASSOCIATED COMPUTATIONAL COMPLEXITY IN NUMBER OF NAND₂ FOR LDA, QDA AND k-NN

# of Lead	LDA				QDA				k-NN			
	Spe(%)	Sen(%)	Acc(%)	NG(log ₁₀)	Spe(%)	Sen(%)	Acc(%)	NG(log ₁₀)	Spe(%)	Sen(%)	Acc(%)	NG(log ₁₀)
1	84.81	64.04	74.42	3.7494	63.08	85.96	74.52	4.1698	78.85	80.38	79.62	5.3833
2	81.54	86.92	84.23	3.8743	68.27	89.23	78.75	4.9507	91.54	84.42	87.98	5.6281
3	93.27	83.65	88.46	4.1754	83.08	82.50	82.79	5.0691	90.19	82.12	86.15	6.1270
4	93.46	84.81	89.13	4.2265	81.92	85.00	83.46	5.2659	90.00	84.23	87.12	6.1826
5	92.88	85.96	89.42	4.1754	80.96	87.69	84.33	5.2659	94.62	85.96	90.29	6.2762

TABLE V
SPECIFICITY, SENSITIVITY, ACCURACY, NUMBER OF SVs, C_s AND ASSOCIATED COMPUTATIONAL COMPLEXITY IN NUMBER OF NAND₂ (SVM CASE 1)

# of Lead	Spe (%)	Sen (%)	Acc (%)	# of SVs	C_s	NG (log ₁₀)
SVM _L 1	61.15	89.04	75.10	86	1	5.8032
2	81.73	88.65	85.19	72	1	5.8857
3	89.81	83.08	86.44	72	1	6.0115
4	85.19	86.54	85.87	64	1	6.0555
5	88.08	85.96	87.02	60	1	6.1056
SVM _Q 1	76.15	86.35	81.25	83	1	5.8700
2	90.77	86.54	88.65	59	1	5.8850
3	92.50	84.42	88.46	50	1	5.9690
4	94.04	86.92	90.48	44	1	5.8830
5	91.92	87.50	89.71	40	1	5.9136

TABLE VI
SPECIFICITY, SENSITIVITY, ACCURACY, NUMBER OF SVs, C_s^{min} AND ASSOCIATED COMPUTATIONAL COMPLEXITY IN NUMBER OF NAND₂ (SVM CASE 2)

# of Lead	Spe (%)	Sen (%)	Acc (%)	# of SVs	C_s^{min}	NG (log ₁₀)
SVM _L 1	60.96	88.65	74.81	86	4	5.8032
2	82.31	88.27	85.29	68	2048	5.8598
3	90.58	81.92	86.25	68	512	5.9856
4	86.15	86.92	86.54	58	64	6.0105
5	87.88	84.23	86.06	50	32	6.0736
SVM _Q 1	77.50	86.15	81.83	81	128	5.8752
2	90.77	85.19	87.98	40	64	5.7473
3	85.00	79.62	82.31	31	16	5.6686
4	89.04	86.35	87.69	25	64	5.6679
5	92.31	86.92	89.62	24	1024	5.6341

is well-known, in the clinical practice that in several cases an abnormal heart-beat may occur in isolation (e.g. ectopic beats) while in principle the patient's overall condition is diagnosed as normal. Therefore if the isolated abnormal heart-beat is chosen for analysis this may result in misclassifying the entire record. As a result, for consistency check, it is always advisable to consider multiple heart-beats per lead for classification.

To begin with, we employ the feature sets and their combinations as derived from the single-beat-scenario classification described previously. However, for the multiple heart-beats case, since the length of the signals in our two ECG databases are different, the total number of heart-beats for each patient, analysed in our study, is ultimately set to 7, as the minimum number of heart-beats available amongst all 104 records is 7. After isolating 7 beats per lead for each record the automated wave boundary detection algorithm used previously is invoked to obtain the boundaries of P, QRS and T waves within each of the heart-beats.

One representative heart-beat per participating lead from each patient is selected and used to train the classifiers according to the feature selection for each lead scenario discussed in Section IV. Since it has been demonstrated that the best classification performance for each classifier is achieved in the $l = 5$ lead scenario, we only consider this and it is expected to give the best accuracy also in the multiple beat-scenario. In total from each record $7 * l$ beats are used and once classifiers are trained, they are applied individually in each one of the 7 sets (7 single-beat-scenarios) of l beats for each patient. To make the final decision on whether the record is normal or abnormal, using the classification results on these 7 single-beat-scenarios, one simple decision-making scheme is applied. Each ECG record is classified into the class, in which the

majority of the 7 single-beat-scenarios is classified. Since the number of heart-beats is odd a tie cannot occur. This decision rule was suggested by expert cardiologists since in the clinical practice also a small number of ECG beats is investigated and based on these, a decision on the patient's condition is made. Obviously continuous CVD monitoring enables the constant evaluation of successive heart-beats and if more heart-beats are classified as abnormal the alarm is triggered. To evaluate the performance of each classifier we used the same metrics as in single-beat classification scenario. The associated results are shown in Table VII.

From Table VII, it can be seen that in terms of overall accuracy LDA performs almost similarly to QDA, while SVM_L, SVM_Q, k-NN underperform. Although, both SVM techniques demonstrate higher sensitivity than both LDA and QDA their specificity is considerably lower. This translates to these methods having high success in classifying abnormal samples correctly, but are prone to misclassify normal records as abnormal. Overall, the total accuracy of the five classifiers reveals once again that LDA outperforms its counterparts when multiple-beats are considered in the 5 lead scenario.

Note that in the multiple-beat-scenario classification all classifiers show reduced accuracy compared to the single-beat based case. This may be because of the presence of redundant information in the multiple heart-beats. Another possible reason could be there is a need for more exploration on how to optimally combine the feature sets resulted from the single-beat-scenario classification for enhancing the accuracy of multiple heart-beat based classification. This exploration will be carried out in the future. However, from the present results it is evident that when considering the trade-off between computational complexity and accuracy, LDA provides the best classifier choice.

TABLE VII
MULTIPLE HEART-BEAT BASED SIMULATION RESULTS FOR EACH CLASSIFIER

	LDA	QDA	SVM _L	SVM _Q	k-NN
Spe(%)	84.31	84.31	43.14	66.67	72.55
Sen(%)	86.96	84.78	91.30	89.13	84.78
Acc(%)	85.57	84.54	65.98	77.32	78.35

VI. HARDWARE ARCHITECTURE DESIGN FOR LDA

As it is evident from the experimental analysis, LDA exhibits the best trade-off between complexity and accuracy for classifying normal and abnormal ECGs. Therefore, a low-energy VLSI solution of LDA-based classification system is decided to be implemented. This investigation will allow us to comment on the implementability of the proposed classifier on resource constrained ambulatory ECG sensors, used in remote CVD monitoring systems. The reason for implementing the system as an ASIC instead of a microprocessor-based or DSP solution has two folds: it is well established that a general-purpose processor or a DSP-based design for any application consumes at least 2-3 orders more power compared to its equivalent ASIC implementation; and an ASIC design is far more easier to be integrated with the ECG sensor, particularly in body-worn sensor networks. As power consumption is the main constraint in our case, it is preferable to choose an ASIC solution. Thus, we aim to design the system using a minimal number of multiplications. The Haar function is selected as the basis function due to its low arithmetic complexity. In particular, we eliminate the square root and division operations present in the Haar transfer functions by combining the DWT coefficients generation and the calculation of spectral energy. The following notation conventions are used: 1T_5 represents the T wave boundaries of lead 1 at level 5, while ${}^2QT_{345}$ indicates the QT waves boundaries of lead 2 at level 3, 4, 5 and so on.

The block diagram of the architecture is illustrated in Fig. 1(a). The proposed architecture is intended to operate on individual single-beat-scenarios. Considering the performance results in Section IV and Section V, the final implementation of the system is based on the 5 lead scenario of the LDA classifier. At first the DWT cD coefficients at decomposition levels 2-5 are computed, using the DWTLV_m blocks in parallel, where m indicates the index of decomposition level. The corresponding DWTLV_m blocks are selectively activated when necessary, depending on which lead is under consideration since the features associated with each lead may belong to any of these decomposition levels according to Table III. By using the appropriate cD coefficients, the corresponding spectral energies are computed in the CoefSelection&Squaring block (CSS). The entire process is done in a lead-by-lead sequence so as to repeatedly take advantage of these functional blocks in our system. We intentionally use a sequential approach here since, according to the clinical specification, classification of normal and abnormal ECGs within a few seconds time is clinically acceptable. Therefore although the architecture could be made faster by parallelizing the functional blocks and dedicating them for each lead, this will eventually result in a detrimental effect on the overall energy consumption. As we will demonstrate in Section VII, with a clock frequency

of 1 KHz the proposed design satisfies the clinical timing requirements, therefore we believe that there is no need in minimizing latency at the expense of power consumption. Once all desired spectral energy features are obtained, the feature vectors are produced in FeatureVectorGenerator block (FVG), followed by the LDA block which simply implements the LDA algorithm and produces the final results. This is then passed to the OutputLabel block in order to output the final label of the input sample.

Two signals - SigStartTime and SigEndTime are used to indicate the start and end time instances of a single heart-beat within which the DWT coefficients are computed. We also consider that the ECG waves boundaries detection is done by some other blocks outside of the present architecture and assume that the corresponding wave boundaries of ECG clinical parameters (1T_5 , 2T_5 , ${}^2QT_{345}$, 3T_5 , 3QRS_2 , ${}^3QT_{35}$, ${}^7QT_{345}$, 8P_5) are available to the system once we have the ECG signal at the input of the present system. CoefStart and CoefEnd provide the CSS block with the signals that indicate the wave boundaries detected externally and are also used for selecting the set of appropriate cD coefficients for the spectral energy computation, with FeaSel indicating which specific spectral energy the CSS block should compute (refer to Table III), as well as the feature vector FVG to generate. Note that FeaSel is a select signal that selects one of the 13 cases of spectral energy considered in this work.

The associated information processing dataflow is depicted in Fig. 1(b). To describe the flow using an example, consider the input ECG signal, Sig, feeds in random samples with a length of 800 for each lead. SigStartTime and SigEndTime are set to 1 and 800 respectively as the time frame of the signal. As the PQRST complexes for each lead are fed sequentially, the DWT coefficients at the selected decomposition levels are computed in parallel and on-the-fly. The coefficient squaring operation, required for computing spectral energy, is done in the period TP (TP1 to TP5) followed by the sequential addition operation to compute the spectral energy features. Once the entire spectral energy computation is finished, AckLabel is asserted in time period TP5, indicating that FinalLabel is available. The process is shown in the upper half of Fig. 1(b). The lower half of Fig. 1(b) explicitly shows the timing details of when exactly each summation of the squared coefficients is done for each lead in TP1 to TP5. These values are stored temporarily into an intermediate register bank and keep on accumulating along the time from TP1 up until TP5. These are used by the FVG block to generate the feature vector. As the LDA block receives the feature vector, the final classification label is produced during TP5.

To process multiple heart-beats in real-time, the system requires an additional memory bank to store the multiple heart-beats. This is because our system is based on single heart-beat classification thus needs to be applied iteratively for labeling the individual single-beats present in the ECG sample. In the following subsection we describe the details of each blocks comprising the overall system.

In the DWTLV_m block, in accordance with Haar DWT filter transfer function, 2^m consecutive data samples are used to compute the detail coefficients cD at each decomposition level

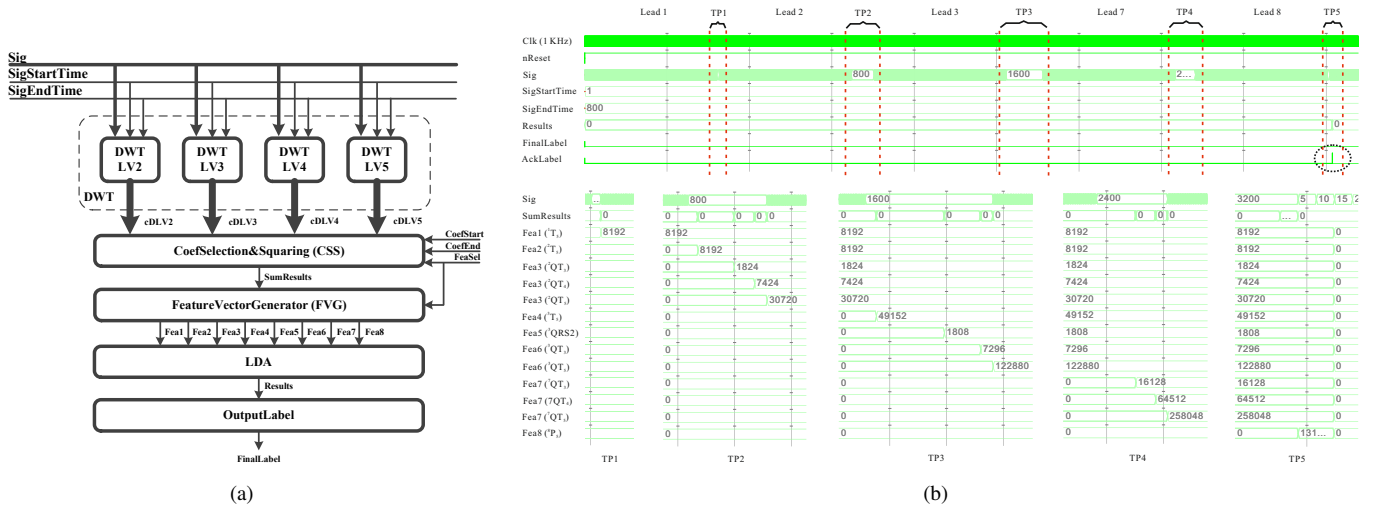


Fig. 1. (a) The hardware architecture of the proposed system; (b) Main data flow of the design

where $m = 2, 3, 4, 5$ in our case. Although the transfer function involves term $1/\sqrt{2}$, it can be ignored in this block and then be considered as shifting in the FVG block, which in turn simplifies the DWT computation block. Therefore, the detail coefficients cD can be computed using simple additions and subtractions. This is done on-the-fly with the incoming signal samples. The architecture for generic coefficient computation at decomposition level m is shown in Fig. 2(a).

The CSS block receives the detail coefficients and produces the spectral energy of a specific ECG clinical parameter at a certain DWT decomposition level. Its block diagram is shown in Fig. 2(b). Four register banks, namely $cDLV2\text{RegBank}$, $cDLV3\text{RegBank}$, $cDLV4\text{RegBank}$, $cDLV5\text{RegBank}$ are used for storing the detail coefficients generated by the $DWTLV_m$ block. Depending on which lead we process, one or more of the register banks are utilized to store the coefficients of certain levels that are necessary in the squaring operation and the others are temporarily ignored during the time of processing that particular lead signal. Once the expected coefficients are successfully stored in the associated register banks, FeaSel is asserted and the corresponding $cD\text{StoreSel}$ signal is generated for selecting the appropriate coefficient register bank(s) for spectral energy computation. A synchronous up-counter is used in 'Squaring' block, with CoefStart and CoefEnd indicating the start and end values of the count operation upon the selected register bank. Henceforth, the coefficients from the CoefStart up to the CoefEnd will be sent to the 'Squaring' sequentially for squaring operation, followed by 'Accumulator' which sequentially sums up squared results. Eventually the overall sum of the selected squared coefficients is outputted as SumResults . Since the timing requirement of our system can tolerate a longer processing time, the CSS block is implemented in such a way that it only requires one multiplier and one adder, with which the squaring and summation of the squared coefficients, thus the spectral energy, are computed.

FVG block is discussed as follows. As previously stated in $DWTLV_m$ block, term $1/\sqrt{2}$ was temporally removed. However, to maintain the computation precision and at the

same time reduce the computational complexity of our system, term $1/\sqrt{2}$ has been implicitly considered in the stage of spectral energy computation. To better explain it in details, Eq.(3) shows the energy calculation procedure of the first squared coefficient at level 2. It can be clearly seen that, at the end of the equation, there is a squared coefficient alongside with the squared term $W_2[0]^2$. This coefficient is resulted from the previous $1/\sqrt{2}$ term in cD coefficient computation. After squaring, ultimately the $1/\sqrt{2}$ term will lead to $1/4$, which can be considered as simple shifting. The same principle can be applied in the decomposition level 3, 4, 5 and thereby reducing arithmetic complexity of the system further.

$$\left[\frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} X[0] + \frac{1}{\sqrt{2}} X[1] \right) - \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} X[2] + \frac{1}{\sqrt{2}} X[3] \right) \right]^2 = \left(\frac{1}{2} \right)^2 [(X[0] + X[1]) - (X[2] + X[3])]^2 = \left(\frac{1}{4} \right) W_2[0]^2 \quad (3)$$

Fig. 2(c) shows the architecture of the FVG block. The above mentioned shifting process is implemented as the first part to adjust the SumResults from the CSS block. Following the shifting, there is a multiplexer associated with $cD\text{StoreSel}$ as the select signal which again is produced from 'Select Signal Convert' with FeaSel as input. This multiplexer is to select the proper shifted SumResults . In order to store it into the InterRegBank at the proper position so that the final features generated as in correct order, 'SelStorePosition' assists to localise the position in InterRegBank according to the FeaSel . Finally, once all the adjusted squared SumResults are stored appropriately, operation of outputting the features will be initiated. Before Fea are sent out, Fea3 , 6 and 7 require one more summation to derive the final result from the associated values in the register bank while the rest are sent straightaway to the LDA block.

Regarding the LDA block, mathematically LDA consists of a linear function coefficients set and one constant, as given in Section II. The coefficients and the constant have already been derived in our simulation (single-beat-scenario classification analysis in Section IV) using Matlab. However, since the constant term turned out to be too small, we ignore it in our LDA

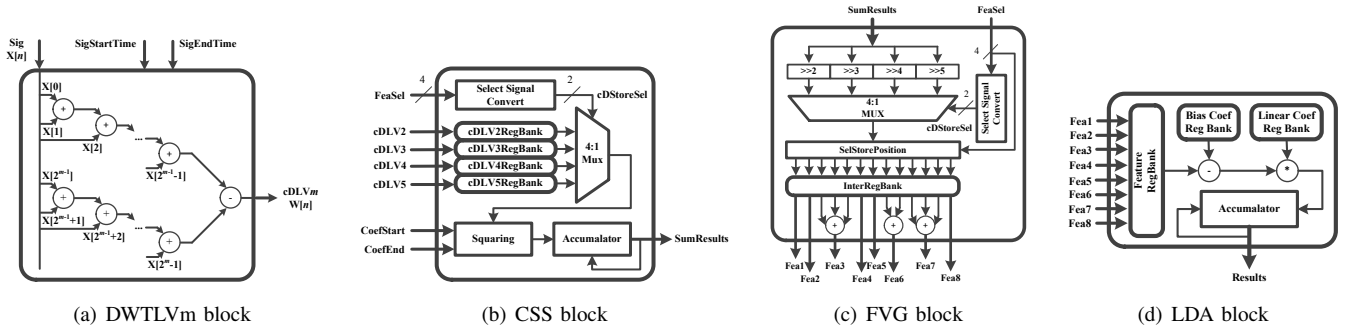


Fig. 2. Hardware implementation of the individual blocks of the proposed architecture

TABLE VIII
SYNTHESIS RESULTS FOR THE PROPOSED SYSTEM

Technology	ST130nm
Global Operating Voltage	1.08V
Clock Frequency	1 KHz
Total Dynamic Power	182.94nW
Total Dynamic Power (LDA block)	17.04nW
Total Cell Area	0.70mm ²
NAND2 Equivalent Area	115.5K

block. Therefore, only LinearCoefRegBank is used to store the set of coefficients. In addition, a normalisation process (i.e. z-score) should also be done on each feature before operating LDA. To simplify this process in our design, the mathematical issues have been considered before the implementation of our design. As we are only concerned about the sign of the results of LDA, the division with the standard deviation can then be eliminated with some expense, and subtraction with associated mean value from each feature is also adjusted accordingly. Henceforth, BiasCoefRegBank is only needed to store these adjusted mean values. With all these values set, Fea1-8 generated from the previous block are firstly subtracted by the corresponding adjusted mean value in sequential order, followed by multiplication with the associated coefficient of the linear function. The result is accumulated with the previous one. When it is done, the final output Results is produced. Fig. 2(d) shows the hardware implementation of the LDA block.

Finally, the OutputLabel block which applies to Fig. 1(a), is used for simply extracting the sign bit of the final output generated from LDA block as the label. This label is categorised into two classes : >0 represents abnormal and <0 normal.

VII. SYSTEM VERIFICATION AND IMPLEMENTATION

The proposed architecture (see Fig. 1(a)) was coded in Verilog and Synopsys Design Compiler was used to synthesize the HDL code at 1 KHz clock frequency and 1.08 V supply voltage, using the STMicroelectronics 130nm technology library. The power consumption of the design was estimated at 182.94 nW, while the LDA block consumed an estimated 17.04 nW using Synopsys PrimeTime. Table VIII illustrates the synthesis results, from where we conclude that the proposed classification architecture is ideal for implementation in low-power mobile CVD platforms and also has the potential to be

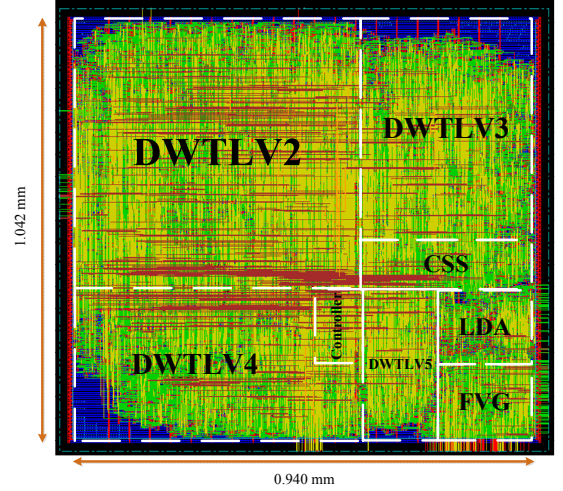


Fig. 3. Core chip layout

integrated with an ambulatory ECG sensor, in the form of a standalone ASIC. Fig. 3 depicts the core layout with labeled blocks after Place and Route (PnR). The associated area is 0.979 mm² and the equivalent NAND2 area is 161.8K. The reason why these measurements are larger than the area in Table VIII is due to practical considerations as cell placing and signal routing are reflected in PnR. The throughput of the design, considering as starting point the time instance when the first ECG sample is fed into the system up to the point when the final label is produced, will vary depending on the number of samples of the heart-beat and also the duration of the ECG clinical parameters, in terms of samples, used in calculating the spectral energy features. To provide an approximation of the throughput, a single-beat-scenario was conducted where the length of the heart-beat has been set to 800 samples and the boundaries of P, QRS, QT, T were set to 128 samples, 96 samples, 416 samples, 160 samples respectively. When considering an ECG sampling frequency of 1KHz (the sampling frequency of the PTBDB) these values are within the normal clinical limits for these parameters. Under this set-up, 4516 clock cycles (at 1KHz operating frequency) are required to output a label for one single-beat-scenario which is approximately 4.5s. Such latency in classifying the ECG signal and to that extent, trigger the “danger” alarm in remote CVD systems is well accepted by cardiologists and

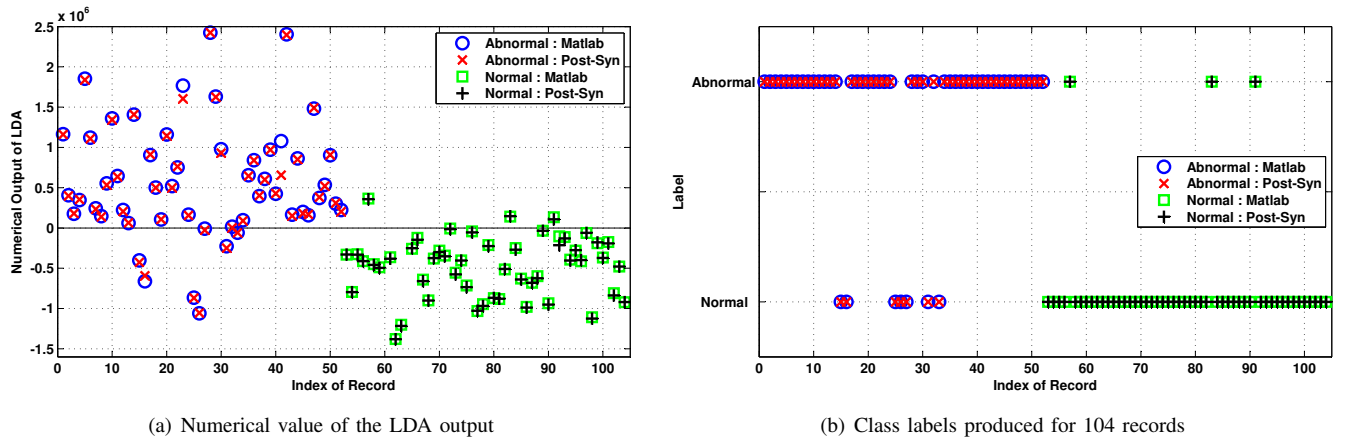


Fig. 4. Comparison between the Matlab and post-synthesis implementations of the proposed classification system

physicians. This fact, in combination with the detrimental effect on the power consumption, prompted us not to consider a parallel architecture in our design.

In addition, a thorough experimental verification was conducted to compare the classification results obtained by the VLSI system at post-synthesis level against the Matlab-based implementation. We considered a single-beat-scenario where 5 leads are available with the LDA as the classification method. The 104 records (52 abnormal and 52 normal), described in Section IV were also utilized for this experimentation. Initially, we used all 104 records in order to train the LDA classifier. For simplicity, this operation took place in Matlab environment. Once the LDA coefficients were derived, they were imported in the synthesised core to define the LDA parameters. In the testing phase the same 104 ECG records were utilized. A Verilog-coded testbench was constructed to define the vectors (104 records) that were used as the input data in the synthesised system. Ultimately, we compared the numerical value of the trained LDA output produced by Matlab-based implementation and the synthesised system. Fig. 4(a) illustrates the the numerical value of the LDA output for both the Matlab-based and the synthesised system for each testing record. It is evident that the values from the two either match or are very close. Fig. 4(b) shows the final classification label for each record based on the operation of the LDA classifier. It can be seen that the classification labels between Matlab and the synthesised system fully agree. The same records were misclassified among the two implementations. In total we observed the same 7 abnormal records misclassified as normal and the same 3 normal misclassified as abnormal out of the total 104, in both implementations. This results in the following values for specificity, sensitivity and overall accuracy, $Spe = 94.23\%$, $Sen = 86.54\%$ and $Acc = 90.38\%$ of the proposed system. This investigation fully validates the synthesised design of the proposed classification system. In essence, also the multiple heart-beat classification scenario is validated here, since this is simply an iterative application of the single-beat method on multiple heart-beats. Finally, due to the fact that the ECG signals considered are actual medical records, we expect the same level of performance in the real-life application of the proposed system.

VIII. CONCLUSION

In this paper, we presented an investigation on the complexity/performance trade-off that various methodologies exhibit in classifying normal and abnormal ECG signals in a remote CVD monitoring system. These systems are viewed as one of the many potential applications of the IoT paradigm. Due to the fact that these systems operate on a limited energy budget our primary focus was to effectively balance performance with computational complexity (and hence energy). Using the spectral energy contained in the constituent ECG waves and calculated using DWT, we first conducted exhaustive simulation experiments considering up to 5 leads to be available, in order to identify the combination of leads (and features) that achieves the best performance in discriminating normal and abnormal ECG signals. In our experiments with 104 ECG records, where both single and multiple-beats classification scenarios were investigated, LDA achieved either the best performance or an accuracy within 7% of the best performing classifier albeit with several orders of less computational demand. Accordingly we presented a low-energy VLSI design of the LDA classifier which consumes only 182.94 nW power when synthesised at 130 nm technology. In addition, we have validated the consistence of our VLSI design, as the classification results obtained from the 104 records through the VLSI implementation completely match the ones produced by the Matlab implementation. The aforementioned findings reveal the potential for ASIC implementation that the proposed solution has, for the classification of normal/abnormal ECGs in resource constrained remote CVD monitoring applications.

REFERENCES

- [1] K. Maharatna *et al.*, "Towards the development of next-generation remote healthcare system: some practical considerations," in *Proc. IEEE Int'l Symposium on Circuits and Systems (ISCAS)*, May 2012, pp. 1–4.
- [2] Frost & Sullivan. (2009, Oct) Preparing for an aging society: challengers faced by healthcare system in European Union, Japan and United States. [Online]. Available: <http://www.frost.com/>
- [3] Z. Loring *et al.*, "A detailed guide for quantification of myocardial scar with the Selvester QRS score in the presence of electrocardiogram confounders," *J. Electrocardiology*, vol. 44, pp. 544–554, May 2011.
- [4] (2012, Dec) Southampton General Hospital. [Online]. Available: www.uhs.nhs.uk/GettingHere/SGH/SouthamptonGeneralHospital.aspx

- [5] (2012, Dec) Policlinico Umberto I - Rome, Italy. [Online]. Available: www.policlinicoumberto1.it/
- [6] (2012, Dec) Chiron Project. [Online]. Available: www.chiron-project.eu
- [7] S. Osowski, T. H. Linh, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 582–589, Apr. 2004.
- [8] F. A. Afsar and M. Arif, "Robust electrocardiogram (ECG) beat classification using discrete wavelet transform," *Physiol. Meas.*, vol. 29, pp. 555–570, May 2008.
- [9] S. Osowski and T. H. Linh, "ECG beat recognition using fuzzy hybrid neural network," *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 1265–1271, Nov. 2001.
- [10] T. H. Linh, S. Osowski, and M. L. Stodolowski, "On-line heart beat recognition using Hermite polynomials and neuron-fuzzy network," *IEEE Trans. Instrum. Meas.*, vol. 52, pp. 1224–1231, Aug. 2003.
- [11] P. de Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of ECG heartbeats using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 1196–1206, Jul. 2004.
- [12] S. Mitra, M. Mitra, and B. B. Chaudhuri, "A rough set-based inference engine for ECG classification," *IEEE Trans. Instrum. Meas.*, vol. 55, pp. 2198–2206, Dec. 2006.
- [13] P. de Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features," *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 2535–2543, Dec. 2006.
- [14] —, "A comparison of the ECG classification performance of different feature sets," in *Computers in Cardiology*, Sep. 2000, pp. 327–330.
- [15] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.
- [16] A. J. Smola and B. Scholkopf, "A tutorial on support vector regression," *Stat. Comput.*, vol. 14, pp. 199–222, Apr. 2004.
- [17] A. Acharyya *et al.*, "Coordinate rotation based low complexity N-D FastICA algorithm and architecture," *IEEE Trans. Signal Process.*, vol. 59, pp. 3997–4011, Apr. 2011.
- [18] B. J. Drew *et al.*, "Comparison of a new reduced lead set ECG with the standard ECG for diagnosing cardiac arrhythmias and myocardial ischemia," *J. Electrocardiology*, vol. 35, pp. 13–21, Jan. 2002.
- [19] E. Mazomenos *et al.*, "A time-domain morphology and gradient based algorithm for ECG feature extraction," in *Industrial Technology, ICIT. IEEE Int. Conf. on*, Mar. 2012, pp. 117–122.
- [20] F. Rincón, J. R. N. Khaled, and D. Atienza, "Development and evaluation of multilead wavelet-based ecg delineation algorithms for embedded wireless sensor nodes," *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, pp. 854–863, Nov. 2011.
- [21] P. S. Addison, *The Illustrated Wavelet Transform Handbook - Introductory Theory and Applications in Science, Engineering, Medicine and Finance*. Bristol, U.K.: Institute of Physics Publishing, 2002.
- [22] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, pp. 215–220, Jun. 2000.
- [23] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Apr. 2003.



Taihai Chen received the B.Eng. degree (1st Class Hons.) from University of Electronic Science and Technology of China (UESTC), Chengdu, China in 2009, and later received the M.Sc. degree (Distinction) from University of Southampton, Southampton, UK in 2010. Since October 2010, he has been working towards his Ph.D degree in Electronic and Software Systems (ESS) research group in University of Southampton.

His research interests include biomedical signal processing, low-power design techniques, pattern recognition, machine learning, associated VLSI architectures, and next-generation healthcare systems.



Evangelos B. Mazomenos received the Diploma degree in electrical and computer engineering from the University of Patras, Patras, Greece, and the Ph.D. degree from the School of Electronics and Computer Science, University of Southampton, Southampton, U.K., in 2006 and 2012, respectively.

Since January 2011, he has been a Research Fellow in the School of Electronics and Computer Science, University of Southampton. His research interests include the area of wireless sensor networks (WSNs) with a focus on positioning and tracking,

biomedical signal processing, and dynamic estimation algorithms.

Dr. Mazomenos is the recipient of the 2009 IET Leslie H. Paddle fellowship on post-graduate studies for his Ph.D. research on real-time target tracking in WSN.



Koushik Maharatna (M'02) received the M.Sc. degree in electronic science from Calcutta University, Calcutta, India, in 1995 and the Ph.D. degree from Jadavpur University, Calcutta, India, in 2002. From 1996 to 2000, he was involved in different projects sponsored by the Government of India undertaken at the Indian Institute of Technology (IIT), Kharagpur, India. From 2000 to 2003, he was a Research Scientist in IHP, Frankfurt (Oder), Germany. During this phase, his main involvement was related to the design of a single-chip modem for the IEEE 802.11a standard. In September 2006, he joined the School of Electronics and Computer Science of the University of Southampton, U.K., where he is currently a Reader. His research interests include low-power VLSI and signal processing for applications in DSP, communication and next-generation healthcare systems, computer arithmetic, analog signal processing, and bio-inspired circuits and systems.

Dr. Maharatna is a member of the IEEE VLSI System Application (VSA) Technical Committee.



Srinandan Dasmahapatra received the B.Sc. degree from Calcutta University, Calcutta, India, in 1986 and the Ph.D. degree in physics from SUNY, Stony Brook, NY, in 1992. He is currently a Lecturer in the School of Electronics and Computer Science, University of Southampton, Southampton, U.K. His research interests include artificial intelligence and pattern recognition, probabilistic knowledge representation, and more recently, systems biology.



Mahesan Niranjan is Professor of Electronics and Computer Science at the University of Southampton, where he was head of the Information: Signals, Images and Systems (ISIS) research group. Prior to this appointment in February 2008, he has held a professorship in the University of Sheffield (1999–2008) and a lectureship in the University of Cambridge (1990–1998). At Sheffield he has served as Head of Computer Science (2002–2004) and Dean of the Faculty of Engineering (2006–2008). He received his BSc from the University of Peradeniya, Sri Lanka

(1982), MEE from Eindhoven, The Netherlands (1985), both in Electronics Engineering, and his PhD from the University of Cambridge (1990). His research interests are in the algorithmic and applied aspects of Machine Learning, and he has authored or co-authored about 100 papers in peer reviewed journals and conferences. He has been Program Chair of several international workshops and has acted as a co-organizer of a six month program on Neural Networks and Machine Learning at the Isaac Newton Institute for Mathematical Sciences, Cambridge. His current research has a strong focus on computational biology and biomedical signal processing.