ON THE TRADE-OFF OF ACCURACY AND COMPUTATIONAL COMPLEXITY FOR CLASSIFYING NORMAL AND ABNORMAL ECG IN REMOTE CVD MONITORING SYSTEMS

T. Chen, E. Mazomenos, K. Maharatna, S. Dasmahapatra and M. Niranjan

School of Electronics and Computer Science, University of Southampton, U.K.
{tc10g09,ebm,km3,sd,mn}@ecs.soton.ac.uk

ABSTRACT

Remote cardiovascular disease monitoring systems are characterised from a limited number of available leads and limited processing capabilities. In this paper, we investigate the trade-off between accuracy and computational complexity in order to derive the best strategy for classifying the ECG signal into normal or abnormal in such systems, with the spectral energy contained in the constituent waves of the ECG signal, as the primary feature for classification. Five established classifiers are considered and through exhaustive simulations the maximum accuracy is derived for each classifier. Based on 104 ECG records, we present a systematic analysis of the trade-off between computational complexity and accuracy, which allow us to deduce the best classification strategy considering only a small number of available leads.

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

1. INTRODUCTION

The recent reports of the World Health Organisation (WHO) state that Cardiovascular Diseases (CVD) remain the biggest cause of death worldwide [1]. Current healthcare systems are characterised by a reactive approach, that delivers care after a critical event has taken place. In recent days, it became clear that by using a proactive approach, it is possible to achieve effective disease management by predicting an impending episode. This concept calls for continuous monitoring of physiological signals. Electrocardiography (ECG) plays an integral role in this approach as it provides an effective screening tool for the evaluation of the patient's condition. Important advancements in Wireless Sensor Networks (WSN) technology enabled the development of portable ECG modules for remote continuous monitoring.

The context of use for these systems is to continuously evaluate the subject's cardiovascular activity and simply create an alarm event whenever an abnormality is detected. For this, an accurate ECG classification mechanism is required. Conventional methods of acquiring and continuously transmitting data to a centralised server for analysis are not applicable in this scenario, because mobile ECG platforms are resource constrained (limited energy supply, processing power) devices. Therefore it is preferable for the signal processing of the acquired ECG signal to take place on the mobile platform

and only the major outcomes (e.g. alarms) will be disseminated. Consequently, in an automated ECG anomaly detection system embedded on a mobile platform, the challenge that rises is the development of an efficient ECG classification strategy balancing accurate performance with the limited processing resources.

Under this context, the aim of this work is to provide a detailed analysis on the trade-off between accuracy and computational complexity for ECG classification. With an eye on mobile healthcare applications, we considered the ECG signal to be available only from a small subset (≤ 5) of the standard 12/15 leads since mobile ECG platforms are equipped only with a limited number of leads, typically 1-5. Automatic classification of ECG signals has been previously reported in the relevant literature and in the majority of approaches time or frequency-domain parameters were chosen as the features for classification [2, 3, 4, 5]. Since the ECG waveform is a direct interpretation of the heart's electrical activity, we hypothesized that potential cardiovascular anomalies would be reflected in the spectral energy of the specific waves that occur during the cardiac cycle. From there, the discriminative features of normal and abnormal ECGs could be derived.

Considering that, remote ECG systems are predominately used to evaluate the subject's overall condition and create an alarm event, but not provide any detailed clinical diagnosis, we only consider two classes in our classification analysis - normal and abnormal ECG. To classify the ECG signal, we investigated five classifiers: linear and quadratic discriminant analysis, support vector machines (SVM) with linear and quadratic kernels and k-nearest neighbours (k-NN). A dataset consisting of 104 ECG records, covering a wide range of heart disease conditions as well as healthy subjects, was used to evaluate the performance of the classifiers. By conducting exhaustive simulations, we are able to designate the combination of leads, from the standard 12 leads, that demonstrates maximum performance for a varying number (1-5) of available leads. The performance investigation is coupled with a detailed complexity analysis of the classifiers considered. In particular for SVM, since their complexity is proportional to the number of support vectors (SVs) employed, we investigated their performance under both the nominal number of support vectors (yields maximum accuracy) and the minimum number of support vectors.

The rest of the paper is structured as follows: in the following section we provide the theoretical background of calculating the spectral energy of the ECG signal. Section 3 describes the process of generating the energy features on a per lead basis, followed by Section 4 which analyses the 5 different classifiers to be used. Section 5 discusses the comparative results obtained from applying the classifiers on the selected database. Finally conclusions are drawn in Section 6.

2. BACKGROUND

The ECG trace is the result of the five composite electrical waves (PQRST) that occur sequentially during a heart cycle (heartbeat). The P-wave is the first deflection to occur in time and corresponds to the depolarization of the atria. Following the P-wave, the QRS-complex is a combination of three successive waves (Q,R,S) which represent the electrical activity of the heart during the depolarization of the ventricles. The T-wave follows the QRS-complex and it is representative of the repolarization of the ventricles. Moreover, the PR-interval (P-onset to QRS-onset) captures the electrical impulse as it is generated in the sinus node up to the point it reaches the ventricles. Similarly the QT-interval (QRS-onset to T-offset) is the part of the ECG that captures all the electrical activity during the ventricular systole [6]. The duration of these ECG parameters (P, QRS, T, PR, QT) is of medical importance and is extensively used in clinical cardiology for the evaluation of the heart's condition.

The approach we followed in this work was to calculate the spectral energy of the ECG signal in the 5 aforementioned temporal parameters and use it as the feature set to base our classification upon. Spectral energy has also been considered for ECG classification in [7], but in their work the authors only investigated the classification of myocardial infarction and normal ECG without taking into account other heart disease conditions. To obtain the spectral energy of these specific time intervals, we use the Discrete Wavelet Transform (DWT) with Haar as the mother wavelet. From the DWT detail coefficients, we obtain not only a measure of the contribution that different frequency components have on the signal, but also their temporal localisation. These properties render DWT ideal for deriving the spectral energy contained within a particular time window of the ECG signal. Since the energy of a wave is predominately contained in the main deflection of the wave, we expect our energy-based features to be more robust against potential inaccuracies in the temporal localisation of the parameters - a typical problem in automated ECG wave localisation algorithms - as long as the main deflection is sufficiently localised and its corresponding DWT detail coefficients are included in the energy calculation.

3. FEATURE GENERATION

To obtain the energy features we initially considered 12-lead ECG signals to be available. One full PQRST-complex was isolated from each lead and the ECG wave boundaries were extracted by utilising our previous work, a Time-Domain Morphology and Gradient (TDMG) based wave localisation algorithm, proposed in [8]. To enhance the level of accuracy of the feature extraction operation and guarantee that the resulting temporal boundaries would include the main deflection of the ECG parameters, we employed the multi-lead wave boundary decision rule proposed in [9], to obtain a global temporal position for each wave boundary (onset, offset). This rule is in essence a k-NN rule, where after the wave boundary estimates of each of the ECG parameters from the 12 leads are sorted from earliest to latest, it designates the first/last estimate with "k" neighbours within a δ ms interval as the global onset/offset of this ECG parameter. In our study we used the same values for "k" and δ as the ones in [9]. However, it must be noted that in a potential implementation of the classification system, the wave boundaries of the ECG parameters will be extracted with feature extraction algorithm on per lead base.

The isolated PQRST-complex was subjected to multi-level Haar DWT decomposition. The ECG intervals, described in Section 2, include parts of the ECG dominated from high frequency components (QRS), and parts that mainly contain low frequency components (P, T-waves) as well as the QT-interval which is characterised by both high and low frequency components. Since different DWT decomposition levels cover different frequency ranges, we chose to analyse the ECG signal in the first 5 levels using Mallat's algorithm. From there the spectral energy of the 5 ECG parameters was calculated using DWT coefficients in different decomposition levels, which is given as

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

where $W_{m,n}$ is the detail coefficient in level m, and n_1 and n_2 denote the onset and offset of the time frame respectively.

Essentially, level 1 was not used as our experiments have shown that this level is dominated by the high-frequency noise components. The QRS-complex energy was calculated using coefficients in level 2 and 3 separately, resulting in two energy features (QRS$_2$, QRS$_3$). The reason for this is that from our experience, QRS components are present both in the frequency ranges covered by these two levels. Moreover, level 5 was employed in the calculation of P, T-wave and PR-interval spectral energy, providing three features (P$_5$, T$_5$, PR$_5$). For the QT-interval energy, which contains both high and low frequency components, we calculated its energy by summing the individual QT-interval 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 energy (QT$_{345}$, QT$_{35}$). The reason for including level 4 in QT$_{345}$ is to capture the transition of the energy from high to low frequency. So, in total 7 distinct energy-based features were 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): QT$_{345}$, QT$_{35}$.

4. CLASSIFICATION METHODS

4.1. Choice of Classifiers

The resource-constrained nature of remote monitoring ECG systems dictates that the employed classification methods must be less computationally demanding in terms of the amount of arithmetic operations required for their implementation. As a result we have selected to investigate Linear/Quadratic Discriminant Analysis as they are considered to be computationally efficient [3]. On the the hand, k-NN [10] and SVM [11] have been previously considered for ECG classification in the relevant literature, therefore we included them in this study. Here, we denote $\{\mathbf{x}, y\}$ as the n-dimension feature vector x_i ($i \in \mathbb{Z} \bigcap [1, n]$) and its label y.

4.1.1. Linear/Quadratic Discriminant Analysis

In Linear Discriminant Analysis (LDA), every class distribution is assumed to be Gaussian and the covariance matrices of the classes are identical. By making this assumption, the classifier becomes linear. Similarly to LDA, Quadratic Discriminant Analysis (QDA) assumes that classes follow Gaussian distributions but with different covariance matrix for each class. The associated decision boundary of each classifier is determined 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_i, L_i and b indicate the coefficients and the interception of the hyperplanes respectively.

4.1.2. SVM with Linear/Quadratic Kernel

Support vector machines is one of the binary classifiers based on maximum margin strategy [12]. A kernel $K(\cdot)$ can be used to determine the separating hyperplane in high dimensional feature space when data are not linearly separable. The output of SVM is determined from the following $y(\mathbf{x}) = sgn(\sum_k \alpha_k y(\mathbf{x}_k) K(\mathbf{x}_k, \mathbf{x}) + b)$, where sgn, k and α_k represent sign function, the number of support vectors and the Lagrangian multipliers respectively. Here, $\mathbf{x}$ indicate the new sample for labelling. For Linear (SVM_L) and Quadratic Kernel (SVM_Q), the associated Kernel functions [12] are $K_L = <\mathbf{x}_k, \mathbf{x}>$ and $K_Q = (<\mathbf{x}_k, \mathbf{x}> +1)^2$, where $<>$ denotes inner product operation.

4.1.3. k-Nearest Neighbours (k-NN)

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

4.2. Computational Complexity

For the purpose of our investigation, we focused entirely on calculating the computational complexity of labelling new data. The computational complexity involved in the training process is not of concern, as this task is considered to be undertaken before the deployment of the mobile system. 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 were also considered. Table 1 shows the total number of arithmetic operations involved in labelling a new sample for the 5 classifiers considered.

Table 1. 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

In Table 1, N, M, S indicate the n-dimensional feature vector, the number of SVs in SVM and the number of training samples respectively. Finally, for simplicity subtractions were considered as additions.

To provide an unified metric for evaluating the overall computational complexity of each classifier and allow its comparison, we opted to use the number of 2-input NAND gates (NAND$_2$), as the measure of complexity, that corresponds to the labelling operation of each classifier. Since there is a variety of hardware implementations for different arithmetic units, we consider a flat unfolded architecture without resource sharing or usage of parallel computing. Firstly, each arithmetic operation is expressed in terms of the number of transistors (T) required, for addition $T_+ = 24b$, for multiplication $T_\times = 30b^2 - 36b$, for squaring $T_{()^2} = 30b^2 - 36b$ and finally for square rooting $T_{\sqrt{}} = 18(\frac{b}{2} + 1)(\frac{b}{2} + 3)$, where b is the word length

[13]. Then, as one NAND$_2$ requires 4 transistors, the arithmetic operation above can be expressed equivalently in terms of the number of NAND$_2$ (G) as $G_+ = 6b$, $G_\times = \frac{15}{2}b^2 - 9b$, $G_{()^2} = \frac{15}{2}b^2 - 9b$ and $G_{\sqrt{}} = \frac{9}{2}(\frac{b}{2} + 1)(\frac{b}{2} + 3)$ respectively. Notice that for the four operations ripple carry adders, conventional array multipliers, squarers (considered as multipliers for simplicity) and non-restoring iterative cellular square rooters were considered respectively [13]. Thus, by calculating the number of NAND$_2$ involved in each arithmetic operation and summing them, we obtain the total number of NAND$_2$ (NG) required to perform labelling.

5. EXPERIMENTAL ANALYSIS

5.1. Database

For the experimental analysis we utilised ECG excerpts from two databases. In total we selected 104 12-lead records divided in two categories (normal/abnormal). For each of these records full clinical diagnosis existed. Specifically, we used PTB database (PTBDB) available in PhysioNet [14]. It contains standard 15-lead ECG recordings at 1KHz covering various disease categories. The 52 records for the normal class (healthy control) were obtained from this database. From the PTBDB we also selected 17 records diagnosed with myocardial infarction covering all available (anterior, inferior, lateral and posterior) subclasses, while 18 records were equally collected from the other six disease classes, namely cardiomyopathy, bundle branch block, dysrhythmia, myocardial hypertrophy, valvular heart disease and myocarditis. To equalise the number of normal and abnormal records, 17 ECG signals from patients with diagnosed myocardial scar were obtained from the Southampton General Hospital Cardiology Department (SGHCD) database. These records are standard 12-lead paper ECG, sampled at 500 Hz which were digitised at a rate of 1000 sample/s with the use of the ECGScan software [15]. Since the 17 records from SGHCD had only 12-leads, we considered the same 12-leads from the PTBDB records.

5.2. Feature Ranking and Space Selection

The set of energy-based features was generated independently for each lead in every record. In order to identify the most discriminating features for classification, we employed Fisher's criterion [16] to select one feature from the L, H and B frequency groups separately that is able to maximally separate the two classes for each lead. Fisher's criterion is expressed as the ratio of the between-class variance to the within-class variance and for the case where only two classes are available, is given as $J = \frac{(m_1 - m_2)^2}{\sigma_1^2 + \sigma_2^2}$, where m and σ^2 denotes the mean and variance of each class [17]. Fisher's criterion was applied on all samples of our two-class database with one feature out of seven for each lead at a time (i.e. one-dimensional feature space). This resulted in a score for each feature per lead. The larger the score is, the more separated the mean values as well as the smaller the overlap of the two classes based on that particular feature will be. Therefore, within each frequency group of each lead, features were ranked according to their respective scores and the top one was selected as the distinctive feature for each feature frequency group. The final selected features for each lead are shown in Table 2. This ranking process reduced the number of necessary features that need to be computed and also reduced the execution time during both training and testing.

Having selected the most discriminative features for each lead, we then conducted exhaustive simulations to identify the combina-

Table 2. 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)
Feature Group	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)
Feature Group	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}

Table 3. 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,LHB),(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)

tion of lead(s), that resulted in the maximum accuracy for each classifier and for a given predefined number of leads. Specifically, if we consider l number of leads, we have C_l^{12} number of possible lead combinations, where C_r^k denotes the combination of order r from k elements. Within each lead combination, at least 1 (at most 3) feature(s) in each separate lead(s) were considered as the feature space for the performance evaluation of each classifier. Considering a mobile healthcare monitoring scenario, we restricted in considering up to 5 available leads in our study, thus $l \in \mathbb{Z} \bigcap [1, 5]$. Here, 10-fold cross validation was applied in this study. Before running cross validation, all the samples of each feature were normalised with respect to their mean and standard deviation (i.e. z-score). For parametric models, the adjustable parameters were optimised using the training data. The regularisation parameter C_s of the SVM classifiers was set to 1. In addition, conventional quadratic programming solving method was selected in the training phase for SVM.

Based on the exhaustive simulation results, the lead combination and its associated feature combination that demonstrated maximum accuracy was selected for every lead scenario[1]. Specifically, for $l = i$ ($i \in \mathbb{Z} \bigcap [1, 4]$), the combination of lead(s) that resulted in maximum accuracy was maintained and the one additional lead that resulted in the most accurate combination for $l = i + 1$ was added.

This selection process allowed us to investigate the improvement in the performance of a classifier when adding extra leads, therefore considering additional "discriminative" features. Table 3 shows the consistent lead combinations that achieved the maximum performance for every classifier and all lead scenarios. The "(Lead, Feature) Combination" column lists which features from Table 2 in each feature frequency group for each lead are used.

5.3. Classification Results Analysis and Discussion

In this section, we present performance results obtained from the 5 classifiers on scenarios where up to 5 leads are considered. To evaluate the classifiers performance we used the testing accuracy defined as:

$$Acc = (TP + TN)/NR \times 100\%. \tag{2}$$

where TP is True Positives, TN is True Negatives and NR is the total number of records. In addition, we list the computational complexity that each classifier requires to label a new sample in terms of NG, using the analysis presented in Section 4.2 and considering a word length of $b = 16$ bit.

Table 4 shows the testing accuracy and complexity of LDA, QDA and k-NN for their corresponding best lead combination in

---

[1]Lead(s) scenario refers to l number of leads under consideration.

different lead scenarios. In general, the accuracy of each classifier increases by adding features from more leads. Similarly NG required also grows, as the dimension of the feature vector is increased. The illustration of the testing accuracy can be found in Fig. 1 and Fig. 3.

Table 4. Accuracy and associated computational complexity in number of NAND$_2$ for LDA, QDA and k-NN

# of Lead	LDA		QDA		k-NN	
	Acc (%)	NG (log_{10})	Acc (%)	NG (log_{10})	Acc (%)	NG (log_{10})
1	75.96	3.7494	76.92	4.1698	78.85	5.3833
2	85.58	3.8743	81.93	4.9507	88.46	5.6281
3	90.38	4.1754	85.58	5.0691	89.42	6.1270
4	92.31	4.2265	88.46	5.2659	91.35	6.1826
5	93.27	4.1754	89.42	5.2659	92.31	6.2762

For SVM, the computational complexity and performance of the classifiers depends on the number of SVs employed by the classifier to discriminate the two classes. The regularisation parameter C_s can be tuned to affect the maximization of the margin, thus the number of support vectors [11]. It could therefore influence both the performance and the computational complexity of SVM classifiers. Consequently, we have decided to explore the trade-off between the accuracy and the number of support vectors employed in SVM classification. As mentioned previously, in order to obtain the maximum testing accuracy that SVM_L and SVM_Q can yield and thus define the best lead (and feature) combination, C_s was initially set to 1. Following, by running grid search on C_s ($C_s = 2^i, i \in \mathbb{Z} \bigcap [-15, 15]$) during the training phase, the C_s^{min} value that resulted in the minimum number of SVs was extracted for every lead scenario.
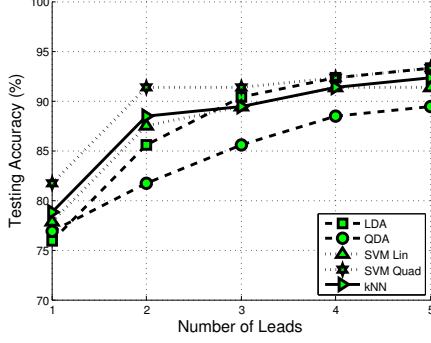
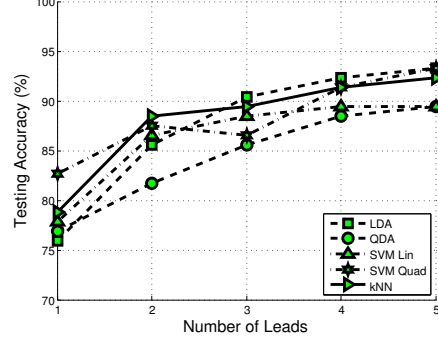**Fig. 1**. Accuracy versus number of leads for all classifiers (for SVM Case 1)

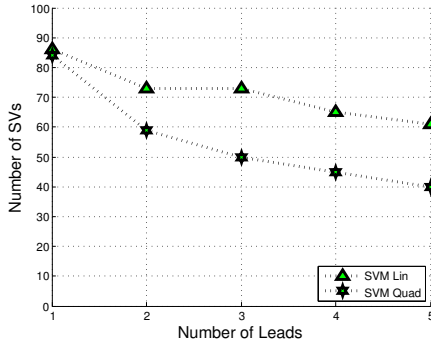**Fig. 3**. Accuracy versus number of leads for all classifiers (for SVM Case 2)

Fig. 2. Number of SVs versus number of leads (for SVM Case 1)
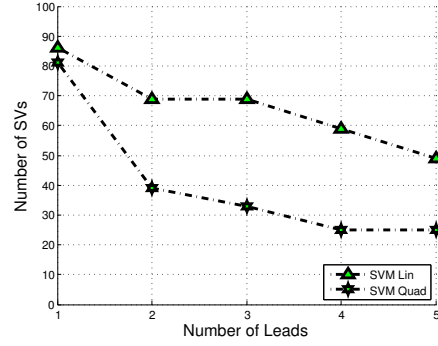

Fig. 4. Number of SVs versus number of leads (for SVM Case 2)

Table 5. Accuracy, number of SVs, C_s and associated computational complexity in number of NAND$_2$ (SVM Case 1)

	# of Lead	Acc (%)	# of SVs	C_s	NG (log_{10})
SVM$_L$	1	77.88	86	1	5.8032
	2	87.50	73	1	5.9100
	3	89.42	73	1	6.0358
	4	91.35	65	1	6.0829
	5	91.35	61	1	6.1349
SVM$_Q$	1	81.93	84	1	5.8910
	2	91.35	59	1	5.8850
	3	91.35	50	1	5.9690
	4	92.31	45	1	5.9232
	5	93.27	40	1	5.9136

Table 6. Accuracy, number of SVs, C_s^{min} and associated computational complexity in number of NAND$_2$ (SVM Case 2)

	# of Lead	Acc (%)	# of SVs	C_s^{min}	NG (log_{10})
SVM$_L$	1	77.88	86	4	5.8032
	2	86.54	69	2048	5.8855
	3	88.46	69	512	6.0114
	4	89.42	59	64	6.0408
	5	89.42	49	32	6.0397
SVM$_Q$	1	82.69	81	128	5.8752
	2	87.50	39	64	5.7052
	3	86.54	33	16	5.7885
	4	91.35	25	64	5.6679
	5	93.27	25	1024	5.7095

We considered two cases in the investigation of the SVM performance. Case 1 is the scenario where we obtain the maximum testing accuracy without being concerned about the number of SVs used; Case 2 is the scenario where we employ the minimum number of SVs and obtain the corresponding accuracy. Performance and complexity results for Case 1 are illustrated in Fig. 1, Fig. 2 and Table 5. As expected, by increasing the number of leads, both SVM$_L$ and SVM$_Q$ exhibit increasing accuracy while the number of required SVs is decreasing. This behaviour was expected since adding more leads increases the size of the feature set, thus maximum accuracy can be achieved with less SVs. Fig. 3, Fig. 4 and Table 6 show SVM results from Case 2. The obtained testing accuracy for SVM$_L$ and SVM$_Q$ indicates that although the performance is in general slightly diminished compared to Case 1, it is still comparable to the ones of

LDA and k-NN while continuing to outperform QDA. The third column lists the minimum number of SVs as obtained by using C_s^{min}. The difference, between Case 1 and Case 2, in the number of SVs for SVM$_L$ is smaller compared to SVM$_Q$, where differences upto 30% are observed for $l > 1$. This significant reduction in the number of SVs leads to a reduction in the computational complexity of the SVM$_Q$ compared to Case 1. By comparing the performance of the 5 classifiers for the different values of l, we observe that SVM$_Q$ for Case 1 shows superior performance in all situations apart from $l = 4, 5$, where LDA achieves the same performance. In addition, LDA outperforms QDA in every occasion while LDA, k-NN and SVM$_L$ achieve comparable performance for $l \leq 3$. Considering the two different cases for SVM, we observe that in Case 2 the total NG are slightly reduced only for SVM$_Q$. Therefore, by combining the

findings in Case 1 and Case 2 for SVM, we argue that the optimisation of the C_s parameter has a positive effect only on SVM_Q in terms of complexity.

By observing the overall NG required from Table 5 and Table 6, 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 increments gradually for LDA, QDA and k-NN while it remains fairly constant for both SVM_L and SVM_Q. The LDA results indicate that while NG is significantly smaller than the other classifiers, the achieved accuracy is either the best (for $l = 4, 5$) or within a 5% margin from the best classifier.

6. CONCLUSION

In this paper we have performed a thorough investigation of the trade-off between performance and computational complexity in the classification of the ECG signal into normal and abnormal. We investigated five classifiers and derived their maximum performance and computational complexity considering 1-5 leads with 104 ECG records. Particularly for SVM, the complexity analysis included both the case of maximum accuracy, without concern of the number of SVs, as well as the case of minimum number of SVs. Our initial findings demonstrated that among the five classifiers, LDA could be a very good choice for the classification of ECG signals in remote CVD monitoring system, effectively balancing accuracy with computational complexity. In the case where only one lead is available, one could consider employing a more complex classifier, e.g. SVM_Q, in order to achieve higher than 80% accuracy. Our future plans will focus on enhancing the classification operation by considering multiple sequential heartbeats from the available ECG leads.

7. ACKNOWLEDGMENT

8. REFERENCES

[1] Frost & Sullivan, "Preparing for an aging society: challengers faced by healthcare system in European Union, Japan and United States," http://www.frost.com/, 2009.

[2] P. D. Chazal, B. G. Celler, and R. B. Reilly, "Using wavelet coefficients for the classification of the electrocardiogram," in *IEEE EMBC*, 2000, pp. 64–67.

[3] P. D. Chazal and R. B. Reilly, "A comparison of the ECG classification performance of different feature sets," in *Computers in Cardiology*, 2000, vol. 27, pp. 327–330.

[4] S.-N. Yu and Y.-H. Chen, "Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network," *Pat. Recog. Letters*, vol. 28, pp. 1142–1150, Jul 2007.

[5] A. S. Al-Fahoum and I. Howitt, "Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias," *Medical & Biological Engineering & Computing*, vol. 37, pp. 566–573, Sep 1999.

[6] N. E. Holtz and T. B. Garcia, *12 lead ECG: the art of interpretation*, Jones and Bartlett Publishers, New York, 2001.

[7] E. S. Jayachandran, P. Joseph K, and R. Acharya U, "Analysis of myocardial infarction using discrete wavelet transform," *Journal of Medical Systems*, vol. 34, pp. 985–992, Dec 2010.

[8] E. B. Mazomenos, T. Chen, A. Acharyya, and *et. al.*, "A time-domain morphology and gradient based algorithm for ECG feature extraction," in *IEEE ICIT*, 2012.

[9] P. Laguna, R. Jan, and P. Caminal, "Automatic detection of wave boundaries in multilead ECG signals: validation with the CSE database," *Computers and Biomedical Research*, vol. 27, pp. 45–60, 1994.

[10] F.-A. A. Minhas and M. Arif, "Robust electrocardiogram (ECG) beat classification using discrete wavelet transform," *Physiological Measurement*, vol. 29, pp. 555–570, May 2008.

[11] S. Osowski, L. T. Hoai, and T. Markiewicz, "Support vector machine-based expert system for reliable heartbeat recognition," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 582–589, Apr 2004.

[12] A. J. Smola and B. Scholkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, pp. 199–222, 2004.

[13] A. Acharyya, K. Maharatna, B. Al-hashimi, and *et. al.*, "Coordinate rotation based low complexity N-D FastICA algorithm and architecture," *IEEE Trans. Signal Process.*, vol. 59, pp. 3997–4011, 2011.

[14] A.L. Goldberger *et. al.*, "PhysioBank, PhysioToolkit, and PhysioNet Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23, 2000.

[15] F. Badilini, T. Erdem, W. Zareba, and *et. al.*, "ECGScan: a method for conversion of paper electrocardiographic printouts to digital electrocardiographic files," *J. Electrocardiology*, vol. 38, no. 4, pp. 310–188, 2005.

[16] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[17] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.