



HAL
open science

Combining deterministic compressed sensing and machine learning for data reduction in connected health

Hassan Rabah, Slavisa Jovanovic, Naeem Ramzan

► **To cite this version:**

Hassan Rabah, Slavisa Jovanovic, Naeem Ramzan. Combining deterministic compressed sensing and machine learning for data reduction in connected health. Muhammad Zeeshan Shakir; Naeem Ramzan. AI for Emerging Verticals: Human-robot computing, sensing and networking, The Institution of Engineering and Technology, 2020, 9781785619823. 10.1049/PBPC034E_ch15 . hal-03981914

HAL Id: hal-03981914

<https://hal.science/hal-03981914v1>

Submitted on 13 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Chapter 1

Combining deterministic compressed sensing and machine learning for data reduction in connected health

*Hassan RABAH¹ Slavisa JOVANOVIC² and
Naeem RAMZAN³*

Connected health is continuously developing, particularly with the advent of the Internet of Things (IoT) interconnecting various sensing nodes to measure a person's vital signs such as electrocardiogram (ECG). Moreover, it is expected an increase in device demand among a currently underserved but significant population. Most of existing devices performing measurement and data transmission require significant effort to integrate more intelligent processing or even decision making at least for data reduction and more autonomy. In this chapter, we propose the combination of a simple compressed sensing measurement technique with a machine learning classification both for data reduction and for low power consumption. The classification is performed on compressed data and the transmission is achieved for warnings only, by sending classification information, in case of a probable pathology detection, and if necessary the compressed data for further analysis. For data acquisition, we utilize a simple deterministic measurement matrix, which facilitates the hardware implementation. The performance of the proposed approach is demonstrated using ECG recordings from three PhysioNet databases: MIT-BIH Arrhythmia Database, MIT-BIH Normal Sinus Rhythm Database, and The BIDMC Congestive Heart Failure Database.

1.1 Introduction

The early and rapid detection of heart disorders is a very important act of surveillance and more critical than treatment. These disorders can be diagnosed by real time analysis of electrocardiogram (ECG) signal. Continuous monitoring of cardiac health is possible and well approved thanks to the wireless body sensor network devices (WBSN) connected to the monitoring center via wireless links [1]. The challenges

¹Institut Jean Lamour, Université de Lorraine

²Institut Jean Lamour, Université de Lorraine

³School of Computing, Engineering and Physical Sciences University of West of Scotland

in such network are mainly : limited available bandwidth, low memory capability, reduced computational power, and limited energy due to the use of small battery [2]. Research has shown that most of the power in a wireless biosensor is consumed by the radio transceiver during the data transmission [3]. Therefore, compressing data can reduce the power consumption of the wireless biosensors during transmission but requires sufficient computation power to perform compression.

Compressed sensing (CS) is a technique suitable for compressing and recovering signals having sparse representations in certain bases [4]. CS has been widely used to compress the data while sensing the signal in the wireless bio-sensors because most of the bio-signals such as ECG have sparse representation in time of a transform domain [5]. The main advantage with CS is that its acquisition process requires less computational power and addresses the constraints of the wireless biosensors. The complexity of the CS encoder depends on the description of the measurement matrix. Initially, random matrices were commonly used. However, they are difficult to implement in hardware. Recently, deterministic measurement matrices have been proposed to facilitate the hardware implementation [6, 7]. On the other hand, the recovery process of CS is computationally complex; and it is generally performed on a distant computer or on the cloud. Thus, even if there exist efficient algorithms and acceleration techniques, it is not suitable for the real time diagnosis.

From the detection point of view, machine learning algorithms for automatic classification of ECG signal has become an increasingly significant topic[8]. In this area, the focus is done on feature selection and extraction [9], and robustness of the machine learning classifiers [10]. In this paper we exploit the combination of compressed sensing and machine learning for data reduction, power consumption and real time detection with respect to some features, instead of a full signal reconstruction. For the CS part we propose the utilization of an encoder based on deterministic measurement matrix of our previous work in association with machine learning to increase the capacity of a WBSN and to reduce the energy consumption of the wireless biosensors. The remainder of this paper is organized as follows. Section 1.2 introduces the basics of CS, the proposed measurement matrix and the principle of classification in compressed domain. Detailed method for compressive sensing, feature extraction and classification in the compressed domain are given in section 1.3. The obtained results are provided in section 1.4 and conclusions are given in section 1.5.

1.2 Background and related work

1.2.1 Compressive sensing

Compressive sensing is a sensing procedure which compresses a signal at the time of acquisition. This technique relies on the sparsity of the signal of interest in the time domain or in a transform basis or in a learned basis. A signal $\mathbf{x} \in \mathbb{R}^N$ is said to be K -sparse or compressible in a basis $\Psi \in \mathbb{R}^{N \times N}$ if its transform $\alpha \in \mathbb{R}^N$ ($\mathbf{x} = \Psi\alpha$) contains at most K non-zero or significant elements such that the remaining $(N - K)$ elements can be discarded without perceptible loss.

The K -sparse signal $\mathbf{x} \in \mathbb{R}^N$ is compressively sensed by multiplying it by a rectangular matrix $\Phi \in \mathbb{R}^{M \times N}$ during the *acquisition process*, where $K < M < N$. The resulting vector $\mathbf{y} \in \mathbb{R}^M$ is called the measurement vector and Φ is called the measurement or sensing matrix. Since \mathbf{x} is sparse in a basis Ψ , \mathbf{y} is expressed as follows:

$$\mathbf{y} = \Phi \mathbf{x} = \Phi \Psi \alpha \quad (1.1)$$

The *recovery process* reconstructs the original signal \mathbf{x} from the measurement vector \mathbf{y} by solving equation (1.1). Since $\Phi \in \mathbb{R}^{M \times N}$ is a rectangular matrix ($M < N$), the problem formulated in equation (1.1) is ill-posed and has infinite solutions. However, based on the knowledge that \mathbf{x} is sparse in a basis Ψ , the recovery process can be performed in two steps.

The first step finds the sparse vector $\tilde{\alpha}$ by solving the following minimization problem:

$$\min_{\tilde{\alpha}} \|\tilde{\alpha}\|_0 \text{ such that } \Phi \Psi \tilde{\alpha} = \mathbf{y} \quad (1.2)$$

Once the vector $\tilde{\alpha}$ has been obtained, the second step reconstructs the original signal as follows:

$$\tilde{\mathbf{x}} = \Psi \tilde{\alpha} \quad (1.3)$$

The appropriate solution to equation (1.2) is generally computed by using convex relaxation (CR) and greedy pursuits (GP) based algorithms such as the orthogonal matching pursuit (OMP). In order to find the unique solution of equation (1.2), the measurement matrix Φ must satisfy the restricted isometry property (RIP) condition [11]. Φ is said to satisfy the K -RIP if there exists a restricted isometry constant (RIC) δ_K , $0 < \delta_K < 1$ such that

$$(1 - \delta_K) \|\alpha\|_2^2 \leq \|\Phi \Psi \alpha\|_2^2 \leq (1 + \delta_K) \|\alpha\|_2^2 \quad (1.4)$$

for all K -sparse vectors α . In general, it is difficult to evaluate the RIP of a given matrix [12, 13]. A sufficient condition used in CS literature is the incoherence between the measurement matrix Φ and the sparsity basis Ψ [14]. The coherence μ between the two matrices is expressed as follows:

$$\mu(\Phi, \Psi) = \max_{i,j} \frac{|\langle \phi_i, \psi_j \rangle|}{\|\phi_i\|_2 \|\psi_j\|_2} \quad (1.5)$$

where $\phi_{i \in \{1, \dots, M\}}$ and $\psi_{j \in \{1, \dots, N\}}$ respectively represent the row vectors of Φ and the column vectors of Ψ . The matrices Φ and Ψ are incoherent if μ is small enough.

Initially random matrices, where the entries are generated by an independent and identically distributed (i.i.d) Gaussian or Bernoulli process, are commonly used since they satisfy the RIP and low coherence with high probability [15]. Recently, more attention has been paid to deterministic measurement matrices [6, 16, 7]. Indeed, the hardware implementation of deterministic measurement matrix is simple. However, the reconstruction of signal $\tilde{\mathbf{x}}$ can be very costly especial the case of IoT with limited computational resources and energy, computing in compressed domain can achieve good results, very reduced data communication, and thus low power.

1.2.2 *Classification in compressed domain*

The last decade had witnessed an important activity in the area of processing in the compressed domain. The first research addressed the problem of efficient feature extraction from compressed video and audio data bases [17]. In the particular case of compressed sensing, where the measurements are obtained by projecting signal in random vectors, it was mathematically proven that CS measurement can be effectively used in signal classification [18]. In [19] Davenport *et al.* demonstrated that small numbers of non-adaptive compressive measurements can suffice to capture the relevant information required for accurate classification. They proposed a *smashed filter* based on matched filter while stressing its compressive nature. Learning directly in the compressed domain has been also demonstrated in [20]. In [21] the authors exploited the discriminative nature of sparse representation to perform classification. Compressed sensing has been combined with support vector machines in [20, 22] and least squares regression [23]; it was shown that training based on the compressed data performs almost as well as the best possible SVM classifier in the data domain. In [24] the authors presented novel approaches that can deal with complex machine learning problems. In particular, they show how to reduce the time needed to train feed-forward neural networks, considering only multilayer perceptrons (MLPs). CS was also associated with deep learning in [25] for images classification. Random Gaussian measurement matrix and discrete cosine transform (DCT) matrix were applied to images and a block-wise histograms used as a feature extractor. A convolutional CS framework was proposed in [26] by associating CS and CNN along with convolutional filters for image classification. Signal processing operations are applied directly to compressively sensed signal for classification by using discrete wavelet transform (DWT) for features extraction and K-means for clustering [27]. In [28] an end-to-end deep learning approach for CL is presented, in which a network composed of fully-connected layers followed by convolutional layers perform the linear sensing and non-linear inference stages. During the training phase, the sensing matrix and the nonlinear inference operator are jointly optimized, and the proposed approach outperforms state-of-the-art for the task of image classification. In [29] the authors investigated the application of compressed sensing in planar tactile arrays. For the measurement matrix, the Scrambled Block Hadamard Ensemble (SBHE) is used. SBHE is a partial block Hadamard transform with randomly permuted columns. The full signal is compressed using the SBHE matrix to generate the compressed signal of a single time instance of contact with an object. Soft-margin SVM is used for classification. Most of the existing works have focused on the image classification applications, and, in the best of our knowledge most of the measurement matrices utilized for CS are random Gaussian. In this work, we propose the association of a simple deterministic measurement matrix associated to SVM classifier for ECG classification in the compressed domain. The following section gives the details of the proposed method.

1.3 Method

The proposed structure for real time ECG monitoring is shown in Figure 1.1. This structure consists of three main operations: Compressive sensing acquisition, features extraction, and classification. In the first operation, the signal $\mathbf{X} = \{\mathbf{x} \in \mathbb{R}^N\}$ is compressively sensed and provides the measurement data $\mathbf{Y} = \{\mathbf{y} \in \mathbb{R}^M\}$. The second operation generates the features \mathbf{F} from the measurements vector. The final operation, after training, classify the features into different classes \mathbf{C} . The details of different operation are given in the following.

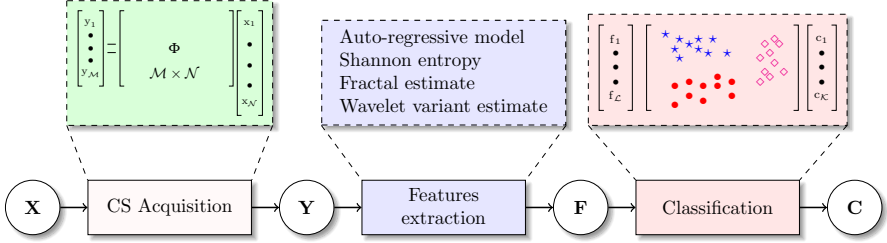


Figure 1.1 General structure of the proposed framework.

1.3.1 Compressive sensing acquisition

In most published papers on classification in compressively sensed domain, the utilized matrices are randomly generated. The complexity of the CS acquisition depends on the description of the measurement matrix Φ . To facilitate the implementation of CS acquisition, particularly for low resources nodes, we utilize a deterministic binary block diagonal (DBBD) matrix developed in our previous work [7], and which is described as follows:

$$\Phi_{\text{DBBD}} = \begin{bmatrix} [1 \dots 1] & 0 & 0 & 0 \\ 0 & [1 \dots 1] & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & [1 \dots 1] \end{bmatrix} \quad (1.6)$$

- Φ_{DBBD} is binary since its elements belong to $\{0, 1\}$ set.
- Φ_{DBBD} is block diagonal.
- Each block, which composes the diagonal of Φ_{DBBD} , has a fixed length $m = \frac{N}{M}$ and can be viewed as a vector $\mathcal{B} \in \mathbb{R}^m$. We set all elements of \mathcal{B} to '1' as shown in equation (1.6).

In order to compare the performance of *DBBD* matrix and the generally used matrix (random matrix), we evaluated the coherence between the inverse DCT (IDCT) matrix and these two matrices versus compression ratio (CR) defined as :

$$CR[\%] = \frac{N - M}{N} \times 100;$$

where N and M are respectively the lengths of the vectors $\mathbf{x} \subset \mathbf{X}$ and $\mathbf{y} \subset \mathbf{Y}$. The results are shown in Figure 1.2.

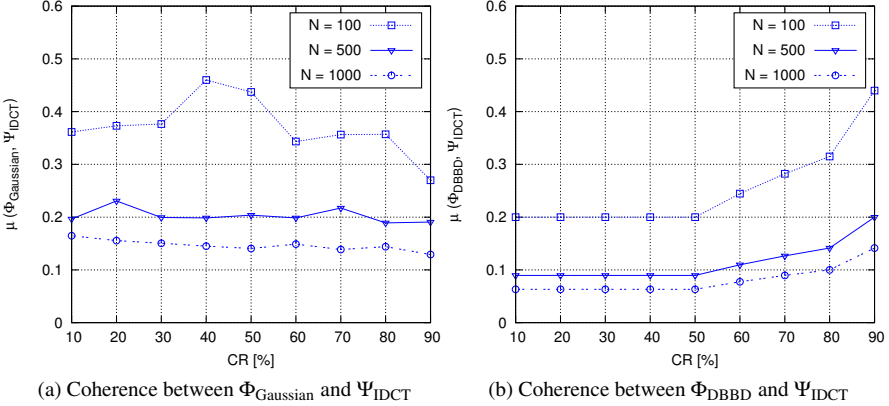


Figure 1.2 *Coherence between measurement matrices and transform domain Ψ_{IDCT} as a function of CR.*

Regardless the measurement matrix, we obtain a lower coherence μ by increasing the value of N . Notice that if CR is less than or equal to 80 %, $\mu(\Phi_{\text{DBBB}}, \Psi_{\text{IDCT}})$ is lower than $\mu(\Phi_{\text{Gaussian}}, \Psi_{\text{IDCT}})$. For a CR of 90 %, we obtain a slightly higher coherence μ with Φ_{DBBB} . However, there is no significant difference between them since the deviation is 0.06, which is much less than the upper boundary of μ .

1.3.2 Features extraction

Feature extraction is an important part of the classification chain. The role of this part is not only reduce data but also to extract relevant information for the classifier. For ECG signal, there is various features that are used to characterise the signal. In this work we will utilise three kernels : auto regressive model, Shannon entropy, and fractal estimates. The feature extraction operation will generate the features vector $\mathbf{F} = \{f_1, f_2, \dots, f_L\}$ consisting of concatenation of the output of the bellow detailed kernels.

Auto-regressive model (AR):

The AR model is applied to obtain the temporal structures of ECG wave-forms. The auto-regressive model of order p consists of representing the signal y_m at a time instant m as a linear combination of p previous values of the same signal. Specifically, the process is modeled as $y_m = \sum_{i=1}^p a_i y_{m-i} + e_m$, where a_i is the i th coefficient of the model and e_m is a white noise with mean zero, and p the AR order. In our study, we will choose $p = 4$ as in [30] where the authors used model order selection methods to determine that a 4th order model provided the best fit for ECG wave-forms; the features are noted $\{ar_1, ar_2, ar_3, ar_4\}$.

Shannon entropy (SE:)

In [31], the authors considered the separation of training and testing samples from the same set of patients. To this end, they proposed a method to classify ECG signals using the Shannon entropy computed on the terminal nodes of a wavelet packet. Shannon wavelet packet entropy is powerful tools for transient signal analysis, which gives information on stroke evaluation. The wavelets are computed down to level 4 using maximal overlap discrete wavelet packet transform, which gives 16 features $\{se_1, se_2, \dots, se_{16}\}$.

Multi-fractal wavelet (MF:)

Alterations to the control system of the heart by myocardial ischaemia are reflected in changes in the complex and irregular fluctuations of the signal. Multifractal analysis is a tool suited for the analysis of this kind of fluctuations, since it gives a description of the singular behavior of a signal. Recently, a new approach for multifractal analysis was proposed, the wavelet leaders, which shows remarkable improvements over previous methods. In [32], the authors propose the application of wavelet leader based multifractal analysis in short-time windows with the aim to characterize and detect ischaemic episodes. Two fractal measures computed by multifractal 1-D wavelet leader estimates are used as features. These features are singularity spectrum, which is as a measure of the multifractal nature of the ECG signal; and the holder exponent describing power-law behavior in the signal at different resolutions. These two features are denoted $\{mf_1, mf_2\}$.

Feature vector:

Compressive sensing is applied to ECG data ($\mathbf{x} \in \mathbb{R}^N$). The compressively sensed data ($\mathbf{y} \in \mathbb{R}^M$) is fed to the feature extraction operation. The ration $R = N/M$ is utilized to define the size of time window for feature extraction. We define T as the total size of non compressed data $\mathbf{X} = \{\mathbf{x}\}$ and T_W time window for features extraction from non compressed data. In this case, the number of features extracted for each window T_W is 22 (4 auto-regressive features, 16 Shannon entropy features and 2 multi-fractal wavelet features). The total number of features for non compressed data is $L = 22 T / T_W$. All this features are concatenated to construct the feature vector $\mathbf{F} = \{ar_1, \dots, ar_4, se_1, \dots, se_{16}, mf_1, mf_2\}$.

When the feature extraction is applied to the compressively sensed data \mathbf{Y} the size of the processing window is reduced with the same factor R as the total data. Thus the number of features remain the same while processing a reduced number of data.

1.3.3 Classification

For the classification of ECG signal in the compressed domain, we make use of Support Vector Machine (SVM), more details can be found for example in [33]. SVM were originally developed for binary classification. For a binary problem, given a set of l data elements \mathbf{x}_i and their corresponding class y_i : $\{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_l, y_l)\}$ where $\mathbf{x}_i \in \mathbb{R}^n$ and $y_i = \pm 1$ the training step consists of resolving the following

quadratic programming problem with linear restrictions (note that here, \mathbf{x} and \mathbf{y} are note the same as in the previous sections).

$$\text{maximize : } \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j), \quad (1.7)$$

$$\text{subject to : } 0 \leq \alpha_i \leq C \quad (i = 1, \dots, l), \sum_{i=1}^l \alpha_i y_i = 0 \quad (1.8)$$

where $\alpha > 0$ are Lagrange multipliers. When the optimization problem is solved, many α_i will be equal to 0, and the others will be Support Vectors. C is positive constant which is chosen empirically by the user. This parameter expresses degree of loosing constraint. A large C can classify training examples more correctly. $K(x_i, x_j)$ is the kernel function. In our case, a multi-class SVM with a quadratic kernel is used.

1.4 Experimental results and discussion

In this section, we present the experimental results for performing the features extraction and classification in the compressively sensed data. We compare obtained results for random measurement matrix and proposed deterministic matrix. We also compare the obtained results for different compressing rates with no compression case.

1.4.1 Datasets

In the numerical experiments, we have used the ECG data from the MIT-BIH Arrhythmia Database corresponding to the normal heartbeat and 2 types of arrhythmia [34, 35]. Each type of heartbeat was extracted from the record which contained most beats of this type. In this work, we utilised the available dataset in [36] where there are 96 recordings from persons with arrhythmia, 30 recordings from persons with congestive heart failure, and 36 recordings from persons with normal sinus rhythms.

1.4.2 Training and validation results

The ECG dataset is described above, it uses 162 measurements sampled at 128 Hz. Each measurement has a size of 2^{16} samples. The measurements are labelled: *ARR* (arrhythmia), *CHF* (congestive heart failure), and *NSR* (normal sinus rhythm). The classifier is trained to distinguish between *ARR*, *CHF*, and *NSR* in two cases: non compressed data and compressively sensed data. In the two cases, 70% of dataset is used for training and the rest for validation. Features are extracted from the two subsets : features for training and features for validation. The feature vectors are composed of 22 features, which are described above. In the case of non compressed data, a time window of size 8192 is utilized to extract 22 features; which gives a total of 176 features for each ECG signal. The same number of features is also extracted from the compressively sensed data. However, the size of time window is divided by the sensing rate thanks to the compression obtained by CS operation.

The procedure is applied to each sensing rate and for the two types of measurement matrices: Random Gaussian matrix Φ_{Gaussian} and deterministic matrix Φ_{DBBD} .

The performance of classifier in classifying the features extracted from original signal as well as from compressive measurements is compared using the standard measures i.e., accuracy, precision, recall (or sensitivity) and F1-score (weighted average of precision and recall). These are calculated from the parameters true positive (TP), true negative (TN), false positive (FP) and false negative (FN) as follows

$$\begin{aligned} \text{Accuracy} &= \frac{TP + TN}{TP + FP + FN + TN} & \text{Precision} &= \frac{TP}{TP + FP} \\ \text{Recall} &= \frac{TP}{TP + FN} & \text{F1-score} &= 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \end{aligned}$$

Figure 1.3 and figure 1.4 show the values of these parameters in the confusion matrices for the deterministic measurement matrix and Gaussian matrix, with sampling rate varying from 1 to 16. The confusion matrices summarize the distributions for elements belonging to a given class (one per row: ARR, CHF or NSR) being assigned a given class label (one per column: ARR, CHF or NSR). The diagonal elements show the number of correct classification for each of the classes. The matrices show that performance is almost stable as R increases for Φ_{DBBD} matrix except for CHF class for $R = 16$, compared to the non compressed data. However, the performance is degrading with increase of R especially for CHF class. This figures show that the proposed matrix performs better than Gaussian matrix.

Table 1.1 compares the accuracies obtained using original signal and the compressively sensed signal for different sensing rates and for the Gaussian and deterministic measurements matrices. The CS method the proposed deterministic matrix Φ_{DBBD} is able to achieve better accuracies than the random Gaussian matrix Φ_{Gaussian} . Precision, Recall, and F1 Score are also used for the evaluation of the evaluation of the classifier. These metrics are shown respectively in figures 1.5, 1.6 and 1.7 demonstrating again that the proposed Φ_{DBBD} matrix is capable of not only preserving information on ECG classes, but also performing better than widely user random measurement matrix.

1.5 Conclusion

Integration of more intelligent processing and decision making at least for data reduction and more autonomy. The Nyquist sampling technique has shown its limitation due to the huge data that can be generated. Compressive sampling has demonstrated to be very efficient in data reductions; but require expensive departed processing, which is not suitable for continuous monitoring and real time decision making in low power and low resources wireless devices. In this paper we presented method based on a combination of deterministic compressed sensing and machine learning for data reduction in connected health. The method is applied to ECG signal classification in the compressed domains without reconstructing the signal. We demonstrate that is possible to capture the difference between the normal heartbeat and 2 types

of arrhythmia by the performance of the SVM classifier on the test set. We also compared the performances of a simple deterministic measurement matrix with random Gaussian measurement matrix. The obtained results, particularly precision and accuracy, of the deterministic matrix are far superior to that of random Gaussian matrix.

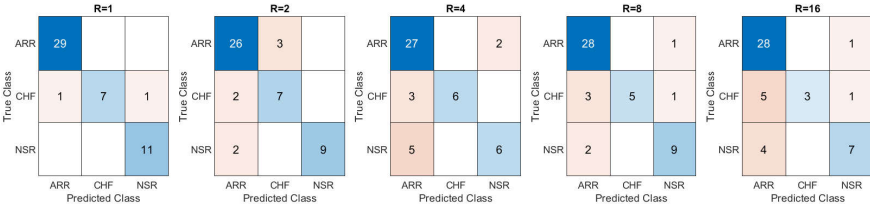


Figure 1.3 *Confusion matrices of classification results by SVM classifier sensing matrix Φ_{DBBD} with sampling rates varying from $R = 1$ to 16*

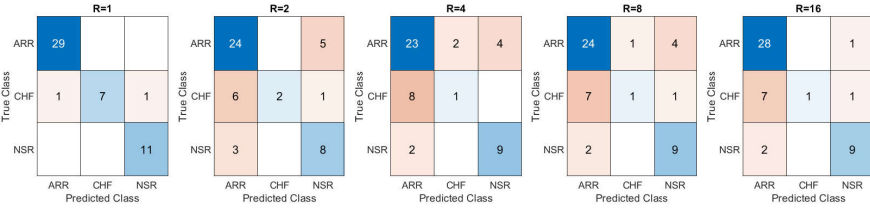


Figure 1.4 *Confusion matrices of classification results by SVM classifier sensing matrix $\Phi_{Gaussian}$ with sampling rates varying from $R = 1$ to 16*

Table 1.1 *Accuracies of SVM applied to the original signal and compressive measurements for Gaussian matrix and deterministic matrix for different sampling rates $R=N/M$.*

Sensing Rate	CS Signal Dimention	$\Phi_{Gaussian}$ accuracy (%)	Φ_{DBBD} accuracy (%)
1 (no CS)	8192	95.91	95.91
2	4096	69.38	85.71
4	2048	67.34	79.59
8	1024	69.38	85.71
16	512	77.55	77.55

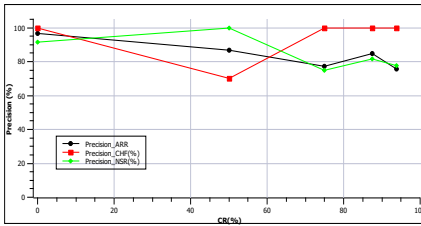
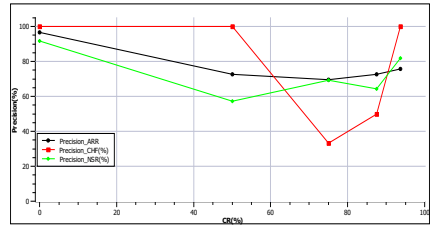
(a) Φ_{DBDD} sensing matrix CR.(b) Φ_{Gaussian} sensing matrix

Figure 1.5 Precision for ARR, CHF, and NSR as a function of CR.

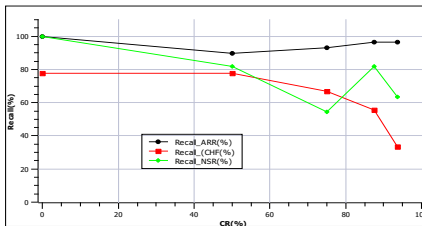
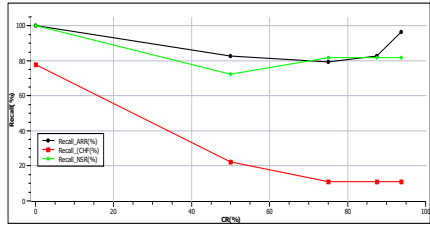
(a) Φ_{DBDD} sensing matrix(b) Φ_{Gaussian} sensing matrix

Figure 1.6 Recall for ARR, CHF, and NSR as a function of CR.

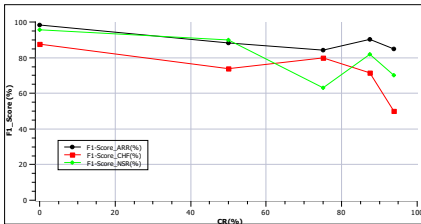
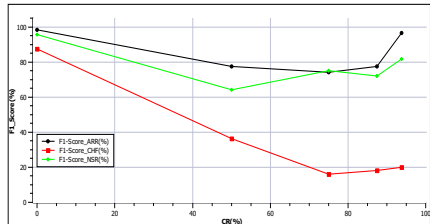
(a) Φ_{DBDD} sensing matrix(b) Φ_{Gaussian} sensing matrix

Figure 1.7 F1-Score for ARR, CHF, and NSR as a function of CR.

References

- [1] Pantelopoulos A, Bourbakis N. A Survey on Wearable Sensor-Based Systems for Health Monitoring and Prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*. 2010 Jan;40(1):1–12. 1
- [2] Pal R, Gupta B, Prasad N, et al. Efficient data processing in ultra low power wireless networks: Ideas from compressed sensing. In: *2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies, 2009. ISABEL 2009; 2009*. p. 1–2. 2

- [3] Chen F, Chandrakasan A, Stojanovic V. Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors. *IEEE Journal of Solid-State Circuits*. 2012;47(3):744–756. 2
- [4] Donoho D. Compressed sensing. *IEEE Transactions on Information Theory*. 2006 Apr;52(4):1289–1306. 2
- [5] Ciocoiu I. Foveated Compressed Sensing. *Circuits, Systems, and Signal Processing*. 2014;p. 1–15. Available from: <http://dx.doi.org/10.1007/s00034-014-9878-2>. 2
- [6] Zeng L, Zhang X, Chen L, et al. Deterministic Construction of Toeplitz Structurally Chaotic Matrix for Compressed Sensing. *Circuits, Systems, and Signal Processing*. 2014;p. 1–17. Available from: <http://dx.doi.org/10.1007/s00034-014-9873-7>. 2, 3
- [7] Ravelomanantsoa A, Rabah H, Rouane A. Compressed Sensing: A Simple Deterministic Measurement Matrix and a Fast Recovery Algorithm. *IEEE Transactions on Instrumentation and Measurement*. 2015 dec;64(12):3405–3413. 2, 3, 5
- [8] Mustaqeem A, Anwar SM, Majid M. Multiclass Classification of Cardiac Arrhythmia Using Improved Feature Selection and SVM Invariants. *Computational and Mathematical Methods in Medicine*. 2018;2018:1–10. 2
- [9] de Chazal P, ODwyer M, Reilly RB. Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features. *IEEE Transactions on Biomedical Engineering*. 2004 jul;51(7):1196–1206. 2
- [10] Ince T, Kiranyaz S, Gabbouj M. A Generic and Robust System for Automated Patient-Specific Classification of ECG Signals. *IEEE Transactions on Biomedical Engineering*. 2009 may;56(5):1415–1426. 2
- [11] Candès EJ. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathématique*. 2008;346(9):589–592. Available from: <http://www.sciencedirect.com/science/article/pii/S1631073X08000964>. 3
- [12] Ben-Romdhane M, Rebai C, Desgreys P, et al. Flexible baseband analog front-end for NUS based multistandard receiver. In: *Joint IEEE North-East Workshop on Circuits and Systems and TAISA Conference, 2009. NEWCAS-TAISA '09; 2009*. p. 1–4. 3
- [13] Bandeira A, Dobriban E, Mixon DG, et al. Certifying the Restricted Isometry Property is Hard. *IEEE Transactions on Information Theory*. 2013 Jun;59(6):3448–3450. 3
- [14] Gangopadhyay D, Allstot E, Dixon A, et al. Compressed Sensing Analog Front-End for Bio-Sensor Applications. *IEEE Journal of Solid-State Circuits*. 2014;49(2):426–438. 3
- [15] Baraniuk R. Compressive Sensing [Lecture Notes]. *IEEE Signal Processing Magazine*. 2007 Jul;24(4):118–121. 3
- [16] Tang Y, Lv G, Yin K. Deterministic Sensing Matrices Based on Multidimensional Pseudo-Random Sequences. *Circuits, Systems, and Signal Processing*. 2014;33(5):1597–1610. Available from: <http://dx.doi.org/10.1007/s00034-013-9701-5>. 3

- [17] Wang H, Divakaran A, Vetro A, et al. Survey of compressed-domain features used in audio-visual indexing and analysis. *Journal of Visual Communication and Image Representation*. 2003 jun;14(2):150–183. 4
- [18] Haupt J, Castro R, Nowak R, et al. Compressive Sampling for Signal Classification. In: 2006 Fortieth Asilomar Conference on Signals, Systems and Computers. IEEE; 2006. . 4
- [19] Davenport MA, Duarte MF, Wakin MB, et al. The smashed filter for compressive classification and target recognition. In: Bouman CA, Miller EL, Pollak I, editors. *Computational Imaging V*. SPIE; 2007. . 4
- [20] Calderbank R, Jafarpour S, Schapiro R. Compressed learning: Universal sparse dimensionality reduction and learning in the measurement domain. Princeton University; 2009. Available from: <http://dsp.rice.edu/files/cs/cl.pdf>. 4
- [21] Wright J, Yang AY, Ganesh A, et al. Robust Face Recognition via Sparse Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2008 feb;31(2):210–227. 4
- [22] Calderbank R, Jafarpour S. Finding needles in compressed haystacks. In: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2012. . 4
- [23] Maillard OA, Munos R. Compressed Least-Squares Regression. In: NIPS 2009. Vancouver, Canada; 2009. Available from: <https://hal.inria.fr/inria-00419210>. 4
- [24] Fabisch A, Kassahun Y, Wöhrle H, et al. Learning in compressed space. *Neural Networks*. 2013 jun;42:83–93. 4
- [25] Gan Y, Zhuo T, He C. Image classification with a deep network model based on compressive sensing. In: 2014 12th International Conference on Signal Processing (ICSP). IEEE; 2014. . 4
- [26] Snoek J, Rippel O, Swersky K, et al. Scalable Bayesian Optimization Using Deep Neural Networks. In: Bach F, Blei D, editors. *Proceedings of the 32nd International Conference on Machine Learning*. vol. 37 of *Proceedings of Machine Learning Research*. Lille, France: PMLR; 2015. p. 2171–2180. Available from: <http://proceedings.mlr.press/v37/snoek15.html>. 4
- [27] Shoaib M, Jha NK, Verma N. Signal Processing With Direct Computations on Compressively Sensed Data. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*. 2015 jan;23(1):30–43. 4
- [28] Adler A, Elad M, Zibulevsky M. Compressed Learning: A Deep Neural Network Approach. *ArXiv*. 2016;abs/1610.09615. 4
- [29] Hollis B, Patterson S, Trinkle J. Compressed Learning for Tactile Object Recognition. *IEEE Robotics and Automation Letters*. 2018 jul;3(3):1616–1623. 4
- [30] Zhao Q, Zhang L. ECG Feature Extraction and Classification Using Wavelet Transform and Support Vector Machines. In: 2005 International Conference on Neural Networks and Brain. IEEE; 2005. . 6
- [31] Li T, Zhou M. ECG Classification Using Wavelet Packet Entropy and Random Forests. *Entropy*. 2016 aug;18(8):285. 7

- [32] Leonarduzzi RF, Schlotthauer G, Torres ME. Wavelet leader based multifractal analysis of heart rate variability during myocardial ischaemia. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. IEEE; 2010. . 7
- [33] Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press; 2000. 7
- [34] Goldberger AL, Amaral LAN, Glass L, et al. PhysioBank, PhysioToolkit, and PhysioNet. Circulation. 2000 jun;101(23). 8
- [35] Moody GB, Mark RG. The impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine. 2001;20:45–50. 8
- [36] MATLAB, editor. Signal Classification Using Wavelet-Based Features and Support Vector Machines; 2019. 8