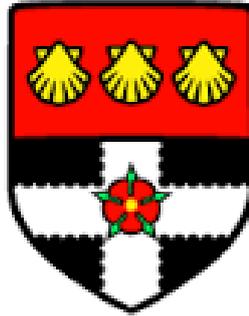


# University of Reading



## ECG Analysis and Classification using CSVM, MSVM and SIMCA Classifiers

A thesis submitted for the degree of Doctor of Philosophy

School of Biological Sciences  
Department of Biomedical Engineering

Najlaa Mohammed Jannah

Supervisors:

1<sup>st</sup> Dr Sillas Hadjiloucas

2<sup>nd</sup> Dr. Faustina Hwang

September 2016

# **Declaration**

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Najlaa Jannah

## Abstract

Reliable ECG classification can potentially lead to better detection methods and increase accurate diagnosis of arrhythmia, thus improving quality of care. This thesis investigated the use of two novel classification algorithms: CSVM and SIMCA, and assessed their performance in classifying ECG beats. The project aimed to introduce a new way to interactively support patient care in and out of the hospital and develop new classification algorithms for arrhythmia detection and diagnosis. Wave (P-QRS-T) detection was performed using the WFDB Software Package and multiresolution wavelets. Fourier and PCs were selected as time-frequency features in the ECG signal; these provided the input to the classifiers in the form of DFT and PCA coefficients. ECG beat classification was performed using binary SVM. MSVM, CSVM, and SIMCA; these were subsequently used for simultaneously classifying either four or six types of cardiac conditions. Binary SVM classification with 100% accuracy was achieved when applied on feature-reduced ECG signals from well-established databases using PCA. The CSVM algorithm and MSVM were used to classify four ECG beat types: NORMAL, PVC, APC, and FUSION or PFUS; these were from the MIT-BIH arrhythmia database (precordial lead group and limb lead II). Different numbers of Fourier coefficients were considered in order to identify the optimal number of features to be presented to the classifier. SMO was used to compute hyper-plane parameters and threshold values for both MSVM and CSVM during the classifier training phase. The best classification accuracy was achieved using fifty Fourier coefficients. With the new CSVM classifier framework, accuracies of 99%, 100%, 98%, and 99% were obtained using datasets from one, two, three, and four precordial leads, respectively. In addition, using CSVM it was possible to successfully classify four types of ECG beat signals extracted from limb lead simultaneously with 97% accuracy, a significant improvement on the 83% accuracy achieved using the MSVM classification model. In addition, further analysis of the following four beat types was made: NORMAL, PVC, SVPB, and FUSION. These signals were obtained from the European ST-T Database. Accuracies between 86% and 94% were obtained for MSVM and CSVM classification, respectively, using 100 Fourier coefficients for reconstructing individual ECG beats. Further analysis presented an effective ECG arrhythmia classification scheme consisting of PCA as a feature reduction method and a SIMCA classifier to differentiate between either four or six different types of arrhythmia. In separate studies, six and four types of beats (including NORMAL, PVC, APC, RBBB, LBBB, and FUSION beats) with time domain features were extracted from the MIT-BIH arrhythmia database and the St Petersburg INCART 12-lead Arrhythmia Database (incartdb) respectively. Between 10 and 30 PCs, coefficients were selected for reconstructing individual ECG beats in the feature selection phase. The average classification accuracy of the proposed scheme was 98.61% and 97.78 % using the limb lead and precordial lead datasets, respectively. In addition, using MSVM and SIMCA classifiers with four ECG beat types achieved an average classification accuracy of 76.83% and 98.33% respectively. The effectiveness of the proposed algorithms was finally confirmed by successfully classifying both the six beat and four beat types of signal respectively with a high accuracy ratio.

# Acknowledgements

All praise and thanks are due to the Almighty Allah who always guides me to the right path and gives me the strength, knowledge, ability and opportunity to undertake this research study and complete this thesis. I would like also to take the time to thank some people who have helped, advised and supported me during my PhD studies.

I would like to express my deepest gratitude to my principal PhD supervisor Dr. Sillas Hadjiloucas for his enthusiasm, invaluable advice and tremendous support throughout the process and the eventual realisation of this dissertation. The amount of help, which he offered me, was incredible and I am very grateful for his generosity. Moreover, I would like to thank my second supervisor Dr. Faustina Hwang who was always a reliable, patient and wise mentor. I would like to extend my gratitude to all of the lecturers who allowed me to attend their lectures during my PhD studies and all members of the School of System Engineering who helped me during my research project.

I wish to express my sincere thanks to Professor Roberto Galvão who has offered invaluable support and advice in developing the research direction for this work and I am very grateful for his huge contribution and continuing support throughout this project. I would like to acknowledge the scholarship support from the Ministry of Higher Education, Kingdom of Saudi Arabia, for its financial support.

I must express my very profound gratitude to my parents for providing me with unflinching support and continuous encouragement throughout my study and writing this thesis, especially my mother (Hayat) for her advice and continuing support. I would also like to extend my thanks to my uncle (Matouq), my brother (Mazen), and my sisters (Lamia, Maram, Mawaheb, Morooj and Mawadah) for their unconditional support during my studies and I would like to tell them that, without their assistance, I would not have been able to complete this project.

Finally, I do not forget to express my deep sense of gratitude to all of my friends. In addition, I would like to express my appreciation and heartfelt thanks to everyone who helped me from my country and from the UK. I am very grateful for their assistance.

## **Dedication**

This thesis is dedicated to my parents for their support and encouragement. Also, it is dedicated to my sisters and brother who have been a great source of motivation and inspiration. I will be forever grateful for their assistance.

# Contents

<b>Declaration .....</b>	<b>ii</b>
<b>Abstract.....</b>	<b>iii</b>
<b>Acknowledgements .....</b>	<b>iv</b>
<b>Dedication .....</b>	<b>v</b>
<b>Contents... ..</b>	<b>vi</b>
<b>List of Figures.....</b>	<b>xii</b>
<b>List of Tables .....</b>	<b>xvi</b>
<b>Abbreviations .....</b>	<b>xviii</b>
<b>Chapter 1 Introduction .....</b>	<b>1</b>
1.1    Background and statement of the research pursued .....	1
1.2    Study aim.....	5
1.3    Hypothesis and objectives of the proposed research .....	6
1.3.1    Hypothesis .....	6
1.3.2    Objectives .....	6
1.4    Contribution of the thesis .....	7
1.5    Organization of the thesis .....	9
<b>Chapter 2 ECG Signal Characterisation Studies and Identification of Arrhythmias in the ECG Signal.....</b>	<b>11</b>
2.1    Introduction .....	11
2.2    The anatomy of the heart and its function .....	11
2.2.1    The heart’s anatomy.....	11
2.2.2    Heart dynamics as related to ECG recordings .....	13
2.3    A concise historical account of work leading to today’s recording practice and interpretation of the ECG signals .....	14
2.3.1    The standard 12-lead ECG record .....	15
2.3.1.1    ECG records using limb leads.....	15
2.3.1.2    ECG records using the precordial leads .....	16
2.3.2    Description of the various ECG wave types recorded .....	17
2.3.2.1    Physiological relevance and good practice in recording the P wave .....	19
2.3.2.2    Typical characteristics of the QRS complex .....	19
2.3.2.3    Typical characteristics of the T wave and its physiological relevance .....	19

2.3.2.4	Typical characteristics of the U wave and its physiological relevance.....	20
2.3.2.5	Typical characteristics of the J wave and its physiological relevance .....	20
2.3.3	The explanation of the physiological relevance of ECG segments and intervals... .....	20
2.3.3.1	Physiological relevance of the PR interval .....	21
2.3.3.2	Physiological relevance of the ST segment .....	21
2.3.3.3	Physiological relevance of the QT interval.....	21
2.3.3.4	Physiological relevance of the PQ interval and ST-T interval.....	21
2.4	Heart rhythms and detection of arrhythmias in the ECG signal.....	22
2.4.1	Sinus node arrhythmias.....	22
2.4.2	Myocardial infarction and ischaemia.....	23
2.4.3	Atrial arrhythmias .....	23
2.4.4	Junctional arrhythmias .....	24
2.4.5	Ventricular arrhythmias .....	25
2.4.6	Atrioventricular blocks .....	26
2.4.7	Bundle branch blocks (BBB).....	27
2.5	Description of the ECG databases used in the current study.....	28
2.5.1	The MIT-BIH Arrhythmia Database .....	28
2.5.2	The European ST-T Database.....	29
2.5.3	The St Petersburg INCART 12-lead arrhythmia database (incartdb).....	30
2.6	Summary.....	31
<b>Chapter 3 Review of ECG analysis and classification algorithms .....</b>		<b>32</b>
3.1	Introduction .....	32
3.2	Pre-processing .....	32
3.2.1	ECG signal denoising .....	33
3.2.2	QRS complex detection .....	35
3.2.3	Recent advances on beat segmentation.....	37
3.3	Feature extraction and selection techniques .....	37
3.4	Literature survey of classification algorithms .....	41
3.4.1	Supervised classification of cardiac arrhythmias.....	42
3.4.1.1	Artificial Neural Networks (ANNs).....	43
3.4.1.1.a	General structure of an neural network .....	43
3.4.1.1.b	RBF classifier .....	43

3.4.1.1.c	Bayesian ANN classifier .....	44
3.4.1.1.d	Extended Kalman Filter (EKF).....	44
3.4.1.1.e	Fuzzy NN.....	45
3.4.1.1.f	Hopfield NN (HNN) classifier .....	46
3.4.1.1.g	Recurrent Neural Network (RNN) and Continuous Time Recurrent Neural Networks (CTRNN) classifiers.....	46
3.4.1.1.h	ECG Classification using other neural network types.....	49
3.4.1.2	Support Vector Machine (SVM) and kernel functions .....	50
3.4.2	Unsupervised classification of cardiac arrhythmias .....	52
3.5	Summary.....	55
<b>Chapter 4</b>	<b>Denoising and classification algorithms using CSVM and SIMCA.....</b>	<b>56</b>
4.1	Introduction .....	56
4.2	Data collection.....	56
4.3	Signal processing software tools used.....	56
4.3.1	MATLAB 2012a software and associated tools box used in the current study..	56
4.3.1.1	Signal processing toolbox .....	57
4.3.1.2	The MATLAB 2012a wavelet toolbox .....	57
4.3.1.3	A Library for Support Vector Machines (LIBSVM) .....	58
4.3.1.4	Wave Form Database (WFDB ) Software Package and toolbox for MATLAB 2012a .....	59
4.4	General classification methodology in separate stages .....	59
4.4.1	Signal pre-processing.....	60
4.4.1.1	Baseline removal and de-noising .....	60
4.4.1.2	R peak detection algorithms in the QRS complex wave and extraction of heartbeats .....	60
4.4.2	Feature extraction and selection .....	65
4.4.2.1	Discrete Fourier Transform (DFT).....	66
4.4.2.2	Discrete Cosine Transform (DCT).....	67
4.4.2.3	Discrete Sine Transform (DST) .....	68
4.4.2.4	Singular value decomposition (SVD) .....	69
4.4.2.5	Principal component analysis (PCA) .....	71
4.4.3	Discussion of the classification algorithm .....	75
4.4.3.1	The necessity for a cross-validation step .....	75

4.4.3.2	Optimization of SVM parameters and training using Sequential Minimal Optimization (SMO) .....	76
4.4.3.2.a	Application of SMO to evaluate the two Lagrange Multipliers in CSVM	77
4.4.3.2.b	Selection procedure for SOM parameters .....	78
4.4.3.2.c	Computing the threshold that defines the CSVM hyperplanes .....	79
4.4.3.3	SVM classification .....	79
4.4.3.3.a	Binary SVM classification.....	80
4.4.3.3.b	Linear SVM classification.....	81
4.4.3.3.c	Soft-margin nonlinear support vector machine .....	83
4.4.3.3.d	Non-linear SVM Classification.....	85
4.4.3.3.e	Types of Kernel Functions used for non-linearly separable datasets .....	87
4.4.3.3.f	Multi-class SVM Classification (MSVM).....	88
4.4.3.3.g	CSVM classification.....	90
4.4.3.4	Soft Independent Modelling of Class Analogy (SIMCA).....	98
4.5	Evaluation of classification results and feature extraction .....	99
4.6	Summary.....	101
<b>Chapter 5 ECG classification studies: algorithmic implementation, results and discussion .....</b>		<b>102</b>
5.1	Introduction .....	102
5.2	ECG beat detection and pre-processing .....	102
5.2.1	Extraction of normal and abnormal heartbeats .....	102
5.3	Pre-processing .....	103
5.4	Features extracted and compressed for each beat.....	104
5.5	Classification studies using the proposed algorithms.....	105
5.5.1	Study one: using binary SVM classification results .....	106
5.5.1.1	Results using the MIT-BIH Arrhythmia Database in datasets from one lead.....	107
5.5.1.2	Results using the MIT-BIH Arrhythmia Database in datasets from two leads...	109
5.5.1.3	Results using MIT-BIH Arrhythmia Database in datasets from two patients ....	110
5.5.1.4	Discussion of SVM binary classification.....	112
5.5.2	Study two: classification using CSVM and MSVM classifiers .....	114
5.5.2.1	Results on MIT-BIH Arrhythmia Database using a precordial lead.....	114

5.5.2.2	Classification of NORMAL, PVC, APC and FUSION beats from the MIT-BIH Arrhythmia Database (using a single limb lead).....	121
5.5.2.3	Results from a MSVM and CSVM study using two leads (V1, V5) with data from the European ST-T database identifying NORMAL, PVC, SVPB and FUSION beats .....	128
5.5.2.4	Discussion of results from the second study using the MSVM and CSVM classifiers .....	131
5.5.3	Study three: SIMCA classifier structure and results.....	135
5.5.3.1	ECG acquisition and pre-processing for the SIMCA study .....	136
5.5.3.2	Feature extraction and calculation of the feature vectors.....	137
5.5.3.3	Classification performance with SIMCA .....	139
5.5.3.4	SIMCA classification using the limb lead datasets from the MIT-BIH Arrhythmia Database .....	140
5.5.3.5	SIMCA Results using the precordial lead dataset for MIT-BIH Arrhythmia records .....	144
5.5.3.6	General discussion of the study classifying six classes of ECG signals .....	148
5.5.4	Study four: Classification structure and results using SIMCA and MSVM classifiers .....	151
5.5.4.1	ECG beat acquisition and pre-processing .....	151
5.5.4.2	Feature extraction.....	152
5.5.4.3	NORMAL, PVC, APC and RBBB beat classification results from the MSVM and SIMCA study using two leads (II, III). .....	152
5.5.4.3.a	MSVM classifier structure.....	152
5.5.4.3.b	SIMCA classifier structure .....	153
5.5.4.3.c	Simulation results .....	153
5.5.4.3.d	Discussion.....	154
5.6	Summary.....	155
<b>Chapter 6</b>	<b>Conclusions and future work .....</b>	<b>157</b>
6.1	Conclusions .....	157
6.2	Limitation of the performed studies .....	160
6.3	Future work using CSVM and SIMCA .....	160
6.4	Trends in mobile computing: using mobile devices and wearable sensors to analyse ECG signals.....	163
6.4.1	Future work on multi-lead analysis using Clifford algebra and Clifford Support Vector Machines.....	165

6.4.1.1	Future work in ECG Multi-lead analysis .....	165
6.4.1.2	Clifford algebra and Clifford Support Vector Machines for multi-lead ECGs... .....	166
<b>References.....</b>		<b>168</b>
<b>Appendix A ECG Signal Processing with Wavelet Filter Banks .....</b>		<b>188</b>
A.1	Introduction .....	188
A.2	Biomedical signal processing and Electrocardiogram (ECG) signals.....	188
A.3	Wavelet transform .....	189
A.3.1	The Continuous Wavelet Transform (CWT) .....	190
A.3.2	The discrete wavelet transform.....	191
A.3.3	Multi-signal Wavelet Analysis .....	192
A.3.4	Mother wavelet choices .....	193
A.3.4.1	The Daubechies wavelets db(N) .....	194
A.3.4.2	The Symlet wavelet (symN).....	195
A.3.4.3	The Coiflet Wavelets (coifN).....	196
A.3.4.4	The Haar wavelet (haar).....	197
A.4	Wavelets from a filter banks perspective .....	197
A.4.1	Wavelet filter bank signal processing of biomedical signals .....	198
A.4.2	Wavelets filter banks and their use in Electrocardiogram (ECG) signals analysis. ... .....	198
A.4.2.1	ECG signal decomposition using Discrete Wavelet Transforms (DWTs)	201
A.4.3	An optimal wavelet filter banks for ECG signal classification .....	204
A.4.4	Wavelet filter bank parametrisation .....	204
A.4.5	Parametrisation and optimisation approach.....	214
A.4.6	Example of a signal decomposition process.....	215
A.5	Hardware requirements in ECG processing .....	216
A.6	Summary.....	217
A.7	References related to the wavelet discussion in the Appendix.....	218
<b>Appendix B Publications presented during the doctoral research.....</b>		<b>221</b>

## List of Figures

Figure 2.1 Anatomy of the human heart, with chambers and valves.....	12
Figure 2.2 The mechanical pumping action of the heart .....	14
Figure 2.3 Standard and augmented limb leads positions .....	16
Figure 2.4 Precordial leads (electrodes) positions .....	17
Figure 2.5 The origin of the electrical activity and electrophysiology of the heart.....	18
Figure 2.6 The basic ECG wave, interval, and segment.....	18
Figure 2.7 The PAC beat were taken from the MIT-BIH Arrhythmia database .....	24
Figure 2.8 The PVC beats were taken from the MIT-BIH Arrhythmia database.....	25
Figure 2.9 The LBBB beats were taken from the MIT-BIH Arrhythmia database .....	27
Figure 2.10 The RBBB beats were taken from the MIT-BIH Arrhythmia database.....	28
Figure 2.11 A segment of an ECG signal taken from the MIT-BIH Arrhythmia database (record no 104).....	29
Figure 2.12 A segment of an ECG signal taken from the European ST-T database (record no e0104) depicting lead III and V4 information .....	30
Figure 3.1 The pre-processing module with methods used in previous studies.....	32
Figure 3.2 ECG feature extraction and dimension reduction techniques .....	38
Figure 3.3 Supervised classification algorithm.....	42
Figure 3.4 General description of the McCulloch and Pitts neurons depicting weights and thresholding function .....	43
Figure 3.5 Elman recurrent neural network structure .....	47
Figure 3.6 Unsupervised classification algorithm. ....	53
Figure 4.1 Procedure of LIBSVM classifiers.....	58
Figure 4.2 Structure and methodology of ECG classification.....	59
Figure 4.3 Cancellation DC Drift and normalisation .....	61
Figure 4.4 ECG Signal after low pass filter (LPF). ....	61
Figure 4.5 ECG Signal after high pass filter (HPF).....	62
Figure 4.6 ECG signal After Derivative. ....	62
Figure 4.7 ECG signal Squaring. ....	62
Figure 4.8 QRS complex Detection. ....	63
Figure 4.9 Reconstruction of detail and approximation coefficients at eight level of decomposition.....	64
Figure 4.10 (a) Typical windowed superimposed ECG datasets from the MIT-BIH cardiac arrhythmia database (record no 104, lead V5, 301 sample window), and (b) time domain filtered and reconstructed signal using 50 Fourier coefficients. ....	67
Figure 4.11 50 Normal ECG beat extracted from ECG datasets from the MIT-BIH arrhythmia database and b) reconstructed signal using 100 Fourier coefficients. ....	68
Figure 4.12 a) Evaluation of matrix effect, b) instrument response function at two wavelengths, c) principal component axes, d) data ranges along the original axes and along the first principal component axis, e) rotation along principal component axes, f) reduction of singular components for PC analysis.....	71

Figure 4.13 10 Normal ECG beat extracted from ECG datasets from the MIT-BIH cardiac arrhythmia database record no 209.....	72
Figure 4.14 The main steps of the PCA algorithm using SVD.....	74
Figure 4.15 Structure of Binary SVM .....	81
Figure 4.16 Structure of soft margin SVM .....	84
Figure 4.17 Transformed data to a higher dimensional space and re-mapped using Radial Basis Kernel function.....	86
Figure 4.18 The structure of multi class SVM using of OVO technique. ....	89
Figure 4.19 The structure of multi class SVM using of OVA technique. ....	90
Figure 4.20 Structure of DAGSVM algorithm .....	90
Figure 4.21 Traditional SVM hyper-plane CSVM. ....	91
Figure 4.22 Complex hyper- plane in CSVM.....	92
Figure 5.1 Normal and abnormal beats extracted from the MIT-BIH database using the accompanying annotation files.....	103
Figure 5.2 Normal beat extracted from record 100 from the MIT-BIH database with (a) elimination of offset , (b) normalized mean-centred data with removed offset. ....	103
Figure 5.3 Block diagram of binary ECG LIBSVM classification.....	106
Figure 5.4 300 Normal and 200 PVC beats were extracted from one lead and patient rec... The time index is 301 points long corresponding to a time duration of 830ms. ....	119
Figure 5.5 LIBSVM classification result from beats extracted from one lead. ....	108
Figure 5.6 ECG beats extract of beats from two leads (a) limb lead (II) and (b) precordial lead (V1).....	109
Figure 5.7 Classification result were beats extracted from two leads. ....	110
Figure 5.8 ECG beats extract from two patients and limb lead (II) (A) patient record no 119 and (B) Patient record no.223. ....	111
Figure 5.9 Classification result for two patients information. ....	111
Figure 5.10 Overview of the pre-processing, feature selection and classification steps associated with the proposed CSVM algorithm used in the multiclass ECG beat classification problem. ....	115
Figure 5.11 Overview of the pre-processing, feature selection and classification steps associated with the proposed SVM algorithm. ....	116
Figure 5.12 Extract of (a) Normal, (b) PVC, (c) APC and (d) PFUS beats.....	118
Figure 5.13 Reconstruction of detail and approximation coefficients at each level of decomposition.....	122
Figure 5.14 Reconstruction using d3, d4 and d5 details coefficients of heart beat. ....	123
Figure 5.15 Detection of the R peak. ....	123
Figure 5.16 Reconstruction of the heartbeat using of d2, d3, d4 and d5 details coefficients. ....	124
Figure 5.17 Detection of the Q peak.....	124
Figure 5.18 Reconstruction wave using d6 and d6 details coefficients.....	124
Figure 5.19 Detection of the T wave peak.....	125
Figure 5.20 Detection of the P peak. ....	125
Figure 5.21 Extract of (A) NORMAL, (B) PVC, (C) APC and (D) FUSION beats .....	125

Figure 5.22 (a) Typical ECG datasets extracted with 301 sample window (b) time domain filtered and reconstructed signal using 50 Fourier coefficients. ....	126
Figure 5.23 (a) Typical ECG datasets extracted with 213 sample window (b) time domain filtered and reconstructed signal using 100 Fourier coefficients. ....	129
Figure 5.24 Block diagram of proposed method for the SIMCA cardiac arrhythmia classification study. ....	137
Figure 5.25 Extract of six types of ECG beats from the MIT-BIH cardiac arrhythmia database. ....	138
Figure 5.26 predicted residual error sum of squares (PRESS) curve for the six beat type classes using ECG beats obtained from the limb lead. ....	141
Figure 5.27 The predicted of input individual ECG beat to the SIMCA classifier ECG beats taking from limb lead (V1) using (a) 10 PCs , (b) 15 PCs, (c) 20 PCs and (d) 30 PCs. ....	144
Figure 5.28 predicted residual error sum of squares (PRESS) curve for the six beat type classes using ECG beats obtained from the precordial lead. ....	145
Figure 5.29 the predicted of input individual ECG beat to the SIMCA classifier ECG beats taking from precordial lead (V1) using (a) 10 PCs , (b) 15 PCs, (c) 20 PCs and (d) 30 PCs. ....	148
Figure A.1: Analysis Normal ECG beat using multi-signal wavelet decomposition (MWD). ....	193
Figure A.2: Analysis Normal ECG beat were taken from St Petersburg INCART 12-lead Arrhythmia Database (incartdb) and extracted coefficients using multi-signal wavelet decomposition (MWD). ....	193
Figure A.3: db1 Scaling and wavelet Function. ....	195
Figure A.4: db2 Scaling and wavelet Function. ....	195
Figure A.5: db4 Scaling and wavelet Function. ....	195
Figure A.6: db6 Scaling and wavelet Function. ....	195
Figure A.7: sym4 scaling and wavelet function. ....	196
Figure A.8: sym6 scaling and wavelet function. ....	196
Figure A.9: coif5 scaling and wavelet Function. ....	196
Figure A.10: Filter bank for perfect reconstruction. ....	197
Figure A.11: Analysis and synthesis filter banks adopted from based on single scale multi-resolution analysis, adopted from . ....	200
Figure A.12: Two-channel filter bank implementation of the wavelet transform applied to a data vector $x$ . Blocks $H$ and $G$ represent a lowpass and a highpass filter respectively and $\downarrow 2$ denotes the operation of dyadic downsampling. The decomposition can be carried out in more resolution levels by successively splitting the lowpass channel, adopted from. ....	202
Figure A.13: ECG signal multilevel decomposition using sym3, db4 and db6 wavelet families ....	204
Figure A.14 Procedure for parameterizing wavelet filter banks by $N$ angles. (a) $S_k$ and $C_k$ represent the sine and cosine of angular parameter $\theta_k$ , respectively ....	208
Figure A.15: Recursive generation of lowpass filter weights $\{h_n^{(k+1)}\}$ in terms of $\{h_n^{(k)}\}$ by adding one additional angular parameter at a time as discussed by Sherlock and Monro. By using this algorithm, any set of $N$ angles $\{\theta_0, \theta_1, \dots, \theta_{N-1}\}$ leads to a sequence of lowpass	

filter weights that satisfies the orthogonality condition were adopted from (A.32), adopted from .....210

Figure A.16: a) Typical signal from the MIT-BIH database with corresponding reconstruction on the basis of approximation and wavelet coefficients at the first decomposition level and b) comparison of filter coefficients impulse response function assuming 12 taps. ....216

Figure A.17: a) Comparison of angular parameters for a standard db6 filter bank with those of a filter bank generated using the proposed procedure and b) normalized frequency response for the two filter banks depicting the difference in their function. ....216

## List of Tables

Table 5.1	Illustration four-dataset groups extracted from MIT-BIH arrhythmia database ...	117
Table 5.2	Classification error rate for different number of DFT coefficients presented to classifier .....	119
Table 5.3	Classification results using DFT extracted coefficients of the ECG signal taken from one lead and CSVM. ....	120
Table 5.4	Classification results using DFT extracted coefficients of the ECG signal taken from two lead and CSVM. ....	120
Table 5.5	Classification results using DFT extracted coefficients of the ECG signal taken from three lead and CSVM. ....	120
Table 5.6	Classification results using DFT extracted coefficients of the ECG signal taken from four lead and CSVM. ....	120
Table 5.7	Collective performance analysis and classification result using DFT and CSVM	121
Table 5.8	Confusion matrix using CSVM .....	127
Table 5.9	Confusion matrix using MSVM .....	127
Table 5.10	Collective performance analysis and classification result using MSVM and CSVM in the study in single lead study presented in section 5.5.2.2.....	127
Table 5.11	Confusion matrix using MSVM .....	130
Table 5.12	Confusion matrix using CSVM .....	130
Table 5.13	Collective PP, ST and SP results using MSVM and CSVM classifiers .....	131
Table 5.14	Illustration of two ECG beat groups from the MIT-BIH cardiac arrhythmia database used in both training and test datasets.....	138
Table 5.15	Classification accuracy rate for different number of PCs coefficients presented to classifier using ECG beats obtain from limb lead (II) .....	141
Table 5.16	Confusion matrix of the ECG signal taken from limb lead and SIMCA classification .....	142
Table 5.17	Collective result performance analysis and classification result using beats obtain from limb lead (II) .....	142
Table 5.18	Confusion matrix of the ECG signal taken from the limb lead using SIMCA classifier and 15 PCs.....	143
Table 5.19	Confusion matrix of the ECG signal taken from the limb lead using SIMCA classifier and 20 PCs.....	143
Table 5.20	Confusion matrix of the ECG signal taken from the limb lead using SIMCA classifier and 30 PCs.....	144
Table 5.21	Classification accuracy rate for different number of PC coefficients presented to the classifier using ECG beats obtained from precordial lead (V1) .....	145
Table 5.22	Confusion matrix of the ECG signal taken from the precordial lead using SIMCA classification .....	146
Table 5.23	Collective result performance analysis and classification result using beats obtain from precordial lead (V1). ....	146

**List of Tables**

---

Table 5.24 Confusion matrix of the ECG signal taken from the precordial lead using SIMCA classification and 10 PCs .....147

Table 5.25 Confusion matrix of the ECG signal taken from the precordial lead using SIMCA classification and 15 PCs .....147

Table 5.26 Confusion matrix of the ECG signal taken from the precordial lead using SIMCA classification and 30 PCs .....147

Table 5.27 Confusion matrix using SIMCA .....154

Table 5.28 Confusion matrix using MSVM .....154

Table 5.29 Collective PP, SE, ACC and SP results using MSVM and SIMCA classifiers...154

Table 5.30 Comparison study four result with other classifiers .....154

Table A.1 Convolution procedure for low-pass filtering showing results before and after dyadic down-sampling.....206

# Abbreviations

ACC	Accuracy
AF	Atrial Flutter
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANNs	Artificial Neural Networks
APC	Atrial Premature Contraction
AR	Autoregressive
ARTMAP	Fuzzy Adaptive Resonance Theory Mapping
BBB	Bundle Branch Blocks
BPNN	Backpropagation Neural Network
HFD	Heart Failure Database
CNNs	Combined Neural Networks
coif	Coiflet Wavelets
CSVM	Complex Support Vector Machine
CR	Compression Ratio
CSVM	Complex Support Vector Machine
CTRNN	Continuous Time Recurrent Neural Networks
CWT	Continuous wavelet transform
DAGSVM	Directed Acyclic Graph Support Vector Machine
db	Daubechies Wavelets
DCTs	Discrete Cosine Transforms
DFT	Discrete Fourier Transform
DOM	Difference Operation Method
DST	Discrete Sine Transforms
DWT	Discrete Wavelet Transforms
ECG	Electrocardiogram
EKF	Extended Kalman Filter
EMD	Empirical Mode Decomposition
EMG	Electromyogram
F	Fusion Beat
PFUS	Fusion of Paced and Normal Beat
FCM	Fuzzy C-means Clustering
FFNNs	Feed Forward Neural Networks
FFT	Fast Fourier Transforms
FTD	Fantasia Database
FUSION	Fusion of Ventricular and Normal Beat
GAs	Genetic Algorithms
GDA	Generalised Discriminant Analysis
harr	Haar wavelet
HCT	Hierarchical Clustering Techniques
HNN	Hopfield Neural Networks classifier
HOS	Higher Order Spectra
HPF	High Pass Filter
ICs	Independent Components
IDWT	Inverse Discrete Wavelet Transform
IPCA	Independent Component Analysis
J	Nodal (junctional) Premature Beat
K-NN	Fuzzy K-nearest Neighbours
LBBB	Left Bundle Branch Block Beat

LDA	Linear Discriminant Analysis
LIBSVM	Library for Support Vector Machines
LPF	Low Pass Filter
LS-SVMs	Fast Least Square- Support Vector Machine
LSVM	Least Square- Support Vector Machine
LVQ	Learning Vector Quantisation
MIT-AD	MIT-BIH Arrhythmia Database
MIT-NSD	MIT-BIH Normal Sinus Rhythm Database
MLPNN	Multilayer Perceptron Neural Network
MSVM	Multi-class Support Vector Machine
MWD	Multi-Signal Wavelet Decomposition
NLMS	Normalised Least-Mean Squares Adaptive Filter
NNs	Neural Networks
NORMAL	Normal Beat
NSR	Normal Sinus Rhythm
ODF	Optimal Decision Function
OSH	Optimal Separation Hyperplane
PB	Paced Beat
PCA	Principal Component Analysis
PCs	Principal Component
PCT	Partitioning Clustering Techniques
PFUS	Fusion of Paced and Normal beat
PLI	Power Line Interface
PNN	Probabilistic Neural Network
PP	Positive Predictively
PR	Perfect Reconstruction
PRD	Root Mean Square Difference
PRESS	Predicted Residual Error Sum of Squares
PVC	Premature Ventricular Contraction
Q	Unknown Beat
QDA	Quadratic Discriminant Analysis
QP	Quadratic Programming
RBBB	Right Bundle Branch Block Beat
RBF	Radial Basis Function
RBK	Gaussian Radial Basis Kernel Function
RLS	Recursive Least Squares
RNN	Recurrent Neural Network
S	Supraventricular Premature Beat
SE	Sensitivity
SGF	Savitzky-Golay Filtering
SIMCA	Soft Independent Modelling of Class Analogy
SMO	Sequential Minimal Optimization
SOCMAC	Self-Organising Cerebellar Model Articulation Controller
SOM	Self-Organizing Maps
SP	Specificity
ST	S-transform
STFT	Short Time Fourier Transforms
SVD	Singular value decomposition
SVM	Support Vector Machine
SVPB	Supraventricular Premature or Ectopic Beats

SVT	Supraventricular Tachycardia.
SVW	Slope Vector Waveform
TWA	T-wave Alternans
VEB	Ventricular Escape Beat
VF	Ventricular Fibrillation
VFW	Ventricular Flutter Wave
VT	Ventricular Tachycardia
WFDB	Wave Form Database
WT	Wavelet transform

# **Chapter 1 Introduction**

## **1.1 Background and statement of the research pursued**

Cardiovascular disease is one of the leading causes of death worldwide. According to public health agencies, cardiovascular diseases are currently one of the biggest causes of death in developed countries and cardiac failure incidents are increasing every year. Early expert diagnoses through Electrocardiogram (ECG) analysis can lead to improved chances of survival for various heart conditions and better management of the disease (in some cases, there is also potential for some reversal of the condition). In addition, ECG analysis is one of the most commonly used methods to establish the onset of heart problems. It can be used for diagnosis of arrhythmias, by providing multi-parametric information regarding both the rhythm and electrical activity of the heart. The ECG signal maps the changes in electrical potential during the cardiac cycle; these are recorded using surface electrodes on the body (arm, leg, and chest wall). The standard ECG contains 12 different leads which record the same electric events but from a different direction. The 12 leads provide a three-dimensional view of the electrical activity of the heart. It can be divided into two groups of six on each limb (the extremity) lead and the chest (precordial) lead [1]. Moreover, the analysis of the electrocardiographic signals from these leads can provide comprehensive information that can be used to classify different heart conditions. The classification of heartbeats on the ECG is important to the study of arrhythmias. In addition, the automation of heartbeat classification could improve the diagnostic quality of arrhythmias, especially in long-term patient's recordings.

Detection and treatment of arrhythmias have become one of the main goals in cardiac care diagnosis provided by general practitioners. In order to prevent a heart attack, one of the most significant ways is to monitor its operation using ECGs. ECG classification of heartbeats enables the identification of specific arrhythmia or other heart conditions. ECG beat classification has been studied extensively using several algorithmic techniques. Some of these methods including analysis of the first derivative of the signal and digital filtering can be used for the ST segment detection and analysis [2].

Of particular interest to the current study are the developed algorithms for automatic ECG heartbeat classification. There are several accounts of ECG beat classification in the literature using different techniques such as neural networks [3], forward neural network [4-5], Recurrent Neural Networks (RNN) [6-7] linear discernment analysis (LDA) [8], probabilistic neural network (PNN) [9-10], self-organizing maps (SOM) [11], deep learning neural network (DNN) [12] and support vector machine (SVM) classifiers [13-14]. Wavelet transforms in conjunction with a novel hybrid neural network have also been discussed in [15]. Multilayer Neural Networks have also been used, especially for classifying the QRS waves [16].

A successful ECG beat classification and arrhythmia detection protocol involve three essential processes: pre-processing, feature extraction and selection, and classification. Feature extraction and dimensionality reduction are important processing steps as they usually affect classifier performance. In the literature on ECG classification, the relative importance of feature selection algorithm has been the subject of considerable discussion. In addition, several procedures have been proposed and developed in order to analyse ECG beats and extract features to be used as input vectors to the classifier. According to Song *et al* (2005) LDA and PCA can be used to obtain four features from the original seventeen features, including two features correlated to rhythm and fifteen features correlated to the morphology of an ECG signal. By combining principal component analysis (PCA) with linear discriminant analysis (LDA), and a PNN classifier, it was possible to categorise eight different types of arrhythmia from ECG beats [17]. A PNN and a backpropagation neural network (BPNN) were the two types of neural networks used with Independent component analysis (IPCA) to differentiate between eight different ECG beat types (normal (NORMAL), Left Bundle Branch Block Beat (LBBB), Right Bundle Branch Block Beat (RBBB), atrial premature contraction (APC), premature ventricular contraction (PVC), paced beat (PB), ventricular flutter wave (VFW), and ventricular escape beat (VEB) ECG beat) with an accuracy above 98%. Using both neural network classifiers with a small number of Independent Components (ICs) as feature vectors provided high classification accuracies [9]. In addition, a combination of various features, such as higher order statistics, morphological features, Fourier transform coefficients, and higher order statistics of the wavelet package coefficients were used to produce an arrhythmia recognition system. These features were used to classify the ECG beats according to five main groups, namely normal beat (NORMAL), ventricular ectopic beat (V), supraventricular ectopic beat (S), fusion beat (F), and unknown beat (Q) using a k-

nearest neighbour classifier. The classification accuracy values were 85.59%, 95.46%, and 99.56% for average sensitivity, selectivity and specificity, respectively [18]. In 2014, Martis *et al.* [19] proposed to use a linear classification method using discrete wavelet transform (DWT) coefficients, featuring also a reduction technique based on the PCA, these were used to extract features from normal and arrhythmia classes. These features were applied to distinguish between normal beats and four arrhythmia classes. A remarkable classification accuracy of 98.78% was shown using a neural network for the classification of the five main beat classes. In another study [20], five types of beats were classified with a lower accuracy of 93.48%. The HOS (higher order spectra) bispectrum after PCA filtering was used to capture features and reduce dimensionality. In addition, a feed-forward NN and a least square-support vector machine (LSVM) classifier were used to classify these five types of beats according to various features. In 2014, Das and Ari used the S-transform (ST) and the wavelet transform to select the features of the ECG beat more effectively. Subsequently, a combination of two features were used to create an input vector. Five classes of ECG beats were identified by using these features as input vectors to a multilayer perceptron neural network (MLPNN) classifier. Average sensitivity performances achieved were 95.70%, 78.05%, 49.60%, 89.68% and 33.89% for NORMAL, S, F, PVC and Q respectively [21]. Lyapunov exponents, wavelet coefficients and power spectral density (PSD) methods were used as a set of features and achieved an average accuracy of 93.89% [22]. The time-domain features, such as morphological and temporal features, were combined to classify the five classes of ECG signal and 85.9% accuracy was obtained using a linear discriminant classifier [23]. In another study, fast least square support vector machines (LS-SVMs) classification with a radial basis function kernel and discrete cosine transform (DCT) were used to classify six types of ECG beats (NORMAL, LBBB, congestive heart failure beat, PVC, non-conducted P-wave, and VEB) obtained from the MIT-BIH database with a 95.2% rate of accuracy [24].

Form the above discussion, it may be concluded that even though many studies have been reported in the arrhythmia beat classification literature, there is a need to improve the classification accuracy for large databases by including different classes and different beats duration. A comparison of beats extracted from the precordial lead or from the limb lead also requires further investigation. In addition, the duration and characteristics of beats may vary according to the type of lead and according to the different types of arrhythmias the segmentation protocol leading to vectors of different dimensions and features. Further studies are therefore needed to develop a holistic and effective treatment programme based on ECG

telemonitoring [25] on the basis of the ECG features identified, and different beat types should be classified using multi-class classification. The development of better diagnostic methodologies could improve the health of people worldwide. The main objective of this thesis is to develop new methodologies for the classification of heartbeats extracted from ECG records.

This project describes the design and implementation of some new ECG beat classification algorithms. This could lead to a new way to support interactive patient care in and outside the hospital and provide services for arrhythmia detection and diagnosis through the use of a new classification technologies. The investigation and evaluation focuses on improving multi-class classification accuracy using Complex Support Vector Machine (CSVM) algorithm [26] to simultaneously classify four types of heartbeats. In addition, the multi-class classification of ECG beats with different beats duration and improve accuracy of classification is also performed. The multi-class classification study is based on the creation of a model for each class independently to classify six types of heartbeats simultaneously using six models instead of one.

The classification performance and generalization ability of the proposed classifier were studied using three publicly available databases the MIT-BIH cardiac arrhythmia database, the European ST-T database and the St Petersburg INCART 12-lead Arrhythmia Database (incartdb). The study focuses on the analysis of both precordial lead and limb beat signals extracted from patient's records.

As already mentioned, pre-processing of the signal and beat detection, feature extraction compression, and classifier implementation are the three most important steps associated with successful ECG arrhythmia classification software. In the first task in this study, ECG beats were extracted from precordial leads and limb leads via base-line removal. Initially, an ECG arrhythmia classification scheme based on PCA feature extraction and binary SVM classification was adopted to discriminate between normal and abnormal beats. Following this, multi-class classifications was investigated using CSVM and Multi-class Support Vector Machine (MSVM). In both CSVM and MSVM algorithms, the ECG arrhythmia classification scheme is based on the use of the Fourier transform for feature extraction, while the CSVM or MSVM classification is adopted to discriminate between four types of cardiac arrhythmias. A DFT algorithm is also used to find the complex Fourier coefficients for each beat presented at the input stages of both classifiers. Thus, frequency-

domain features were also used to create an input vector for the classifier. Finally, both classifier methods are employed to simultaneously classify four types of ECG arrhythmias: NORMAL, PVC, APC and Fusion of ventricular and normal beat (FUSION) or fusion of paced and normal beat (PFUS).

Following on the success of this study, an ECG arrhythmia classification scheme based on a PCA feature extraction scheme and a SIMCA classifier were used to differentiate between six types of arrhythmia conditions. PCA was used to find feature vectors from each beat and present them at the input stages of the Soft Independent Modelling of Class Analogy (SIMCA) classifier. PCA is a statistical method that aims to extract the underlying components from multidimensional data while in its domain. Finally, the SIMCA classifier was employed to create a model for each class individually and to classify six types of ECG beats simultaneously, for NORMAL, PVC, APC, LBBB, RBBB and FUSION beats that were extracted from both precordial and limb leads. The datasets in this study had a variable dimension for the input vector to the classifier. In addition, the final study focuses on the use of a new ECG arrhythmia classification scheme based on PCA for feature extraction and either an MSVM or a SIMCA classifier to differentiate between four type of arrhythmias conditions that were obtained from records in the St Petersburg INCART 12-lead Arrhythmia Database (incartdb) [27]. PCA is used to find the principal components (PCs) associated with each beat and to create feature vectors based on the PCs that were presented at the input stage of the MSVM and SIMCA classifiers. Both the MSVM as well as the SIMCA classifier were used to classify the four types of ECG beats (annotated as NORMAL, PVC, and APC and RBBB respectively) simultaneously.

As a general remark, this study focused on ECG classification using only supervised classification methods. A brief description of unsupervised classification techniques that have been used in ECG beat classification is summarized in chapter three.

## **1.2 Study aim**

The project aims to introduce a new way to support interactively patient care inside and outside the hospital and provide services for arrhythmia detection and diagnosis through the use of new classification algorithms. A supervised methodology for analysing ECG beat signals that were extracted from Holter recordings, including pre-processing, feature selection and classification, was designed to diagnose cardiac arrhythmia conditions. The proposed

methodology will be able to separate four or six types of arrhythmia groups. Thus, the natural question is how we could use the development of traditional SVM algorithms with an extension to complex spaces, or create an individual PC model for each class to classify four and six types of heartbeats, respectively.

## **1.3 Hypothesis and objectives of the proposed research**

### **1.3.1 Hypothesis**

It is hypothesized that the application of CSVM and SIMCA to the classification of individual ECG beat waveforms will provide a better classification accuracy than that other machine-learning and neural-network approach have achieved. In addition, by creating the input vector to the classifier using a selected number of features it is expected this will lead to high classification accuracies and improve the generalization ability of the classifier improving and its accuracy.

### **1.3.2 Objectives**

The following objectives were set at the onset of this study.

1. To firstly make a survey of current literature on ECG signal processing and classification.
2. To design and implement sound procedure for extracting ECG beats based on the Wave Form Database (WFDB) Software Package for R (peak) localization and detect peaks in other waves such as P and T. A further aim to review the different method for QRS complex estimation and identify the main significant characteristics of these waves in relation to signal amplitude and time of deflection.
3. To develop feature extraction and selection algorithms taking into account that these features may to be extracted in either time or frequency domains and assess their effectiveness. A further aim was to identify the best number of features.
4. To develop classification schemes for ECG classifiers based on reduced features in the ECG beats in binary SVM, MSVM, CSVM and SIMCA classifiers.

5. To investigate if signal pre-processing in the wavelet domain can improve classification accuracy in SVM classification and extend SVM classification to higher dimensions to take advantage of developments in future ECG imaging techniques.
6. To assess the performance of the new classification algorithm on the basis of sensitivity (SE), specificity (SP), positive predictively (PP) and accuracy (ACC).

## **1.4 Contribution of the thesis**

The major contribution of this thesis is the development of a novel classification framework for ECG analysis based on two types of algorithms, CSVM and SIMCA. These algorithms are used after compressing the original time domain signals using DFT and PCA. The work systematically changes the number of extracted features presented to the developed classifiers so that an optimal number of features can be identified for each algorithm enabling their performance to be optimized. The work is also one of the first of its kind with respect to multi-lead detection, where several arrhythmia conditions may be simultaneously present in the ECG records that need to be classified.

In chapter two, following a survey of the different heart condition as discussed in the literature, different feature in the time-domains signal are related to specific pathological condition. In addition, details associated with these recordings as found in the most widely used ECG databases are identified.

Chapter 3 provides a comprehensive literature review of existing supervised and unsupervised classification algorithms and discusses different pre-processing algorithms as well. The work contributes to the current understanding of the different algorithms that may be used to perform ECG analysis.

Chapter 4 provides a systematic overview of signal de-noising, compression based on the sine, and cosine transforms as well as PCA. In addition, it discusses the sequential minimal optimization algorithm, which is adopted in the SVM classifier used in the thesis. In addition, it discusses a new extension for CSVM, which has never been used in the context of ECG signal classification. An evaluation procedure, which may be used with the newly developed ECG classifiers, is also adopted.

In chapter five, the results from several studies using the different algorithms (binary SVM, MSVM, CSVM and SIMCA) are presented and discussed. In the first study, single

leads records from a single patient are classified using binary SVM. The 2-lead records from a single patient and one lead records from two patients are classified (binary SVM classification). The second study adopts the CSVM methodology, using either a precordial-lead or a limb lead, assuming Fourier transform pre-processing and feature extraction of the time-domain datasets. Peak detection using the DWT is also performed. ECG beat characteristics were extracted using either the pre-existing annotation files in the adopted database or multiresolution wavelet. Implementation of CSVM classification using frequency domain features (Fourier coefficients) of four types of ECG signal beats showed an improvement in classification accuracy. This is a useful result to the ECG community. Using the ECG beat from two correlated leads also showed a significant improvement in the classification accuracy. This study demonstrates the importance of using many patient records and two leads. This is also the first study of its kind that investigates the application of CSVM to ECG beat classification. The CSVM algorithm should find additional application domains in the biomedical community. Several (four) types of arrhythmia condition are simultaneously classified (multi-class classification) successfully using the proposed method. The third study focuses on multi-class classification using SIMCA with PCA noise reduction. Feature extraction results of the proposed algorithm were compared with other methods in the literature that were used to simultaneously classify six types of ECG beats. This study demonstrates that using a multi-model approach after optimizing their number of PCA coefficients provides better classification accuracy than using a single-model. This is the first study reporting an advantage of creating a model for each class, as well as using different dimension vectors at the input to the classifier. The results showed an improvement in classification accuracy. Furthermore, the algorithm showed potential for multi-class studies where a large number of ECG beat types are present.

The four studies display the distinct advantages in using an extension of SVM to complex spaces as well as SIMCA to perform multi-class classification of multiple ECG recodes simultaneously. In addition, this research confirmed that the use of a selected number of Fourier coefficients or PCs to approximate the ECG beat signal and compress the input features to the classifier could lead to high classification accuracies and improve the generalization ability of the classifier. This project will fill a gap in the literature in developing classification algorithms that can be used to classify heart problems and diagnose cardiac arrhythmia conditions with higher classification accuracy.

## 1.5 Organization of the thesis

This thesis is structured around five main topics: *introduction, a literature review on ECG classification algorithms, the provision of new investigation using the various algorithms, the analysis of results and their evaluation, and finally the generation of conclusions and the provision of directions for future work.* The chapters are organised as follows:

*Chapter 1* gives a brief research background about the problem and discusses the methods adopted to do the research. The aims and objectives of this research are presented in this chapter as well.

*Chapter 2* presents a literature review of the current understanding of the anatomy of the human heart and its working, the measurement of ECGs and the various ECG interpretations based on identified waves. The basic modes of ECG beat recording and ECG wave morphology are briefly described according to different modes of data acquisition on the basis of lead placement. In addition, some types of arrhythmias and details of the three databases that were used in this study are explained briefly.

*Chapter 3* focuses on the literature review of some previous studies related to ECG analysis and classification algorithms. This chapter also provides the theoretical background to ECG classification, including the methods used for feature selection and classification according to previous studies found in the literature. Supervised and unsupervised classification schemes are also mentioned. The use of SVM and neural networks in ECG classification are discussed.

*Chapter 4* gives a background to the current methods, used by physicians to diagnose arrhythmia conditions. It also discusses the theoretical and technical background of the algorithms that were used to develop the feature selection and the classification scheme. In addition, it describes some MATLAB 2012 a toolboxes that were used for the pre-processing and classification of ECG beats. The signal pre-processing and feature estimation techniques are also described. Furthermore, this chapter provides a concise description of Fourier transform methods such as DFT, DST, DCT, and SMO algorithms, as well as an introduction to CSVM and kernel functions. Moreover, it provides a review of the PCA and SIMCA classifier algorithms. It finally describes the methods and tools used to evaluate the classification results.

*Chapter 5* describes all the experimental protocols and procedures, as well as shows the results of the investigation and provides their discussion. Each section explains how the experiments were conducted and how the results were obtained. The chapter presents in more detail the actual algorithmic implementation and performance evaluations of the all the proposed algorithm and classification results. Furthermore, it discusses ECG beat extraction via the pre-processing and feature extraction algorithms, focusing on the results using the various methods, such as DWT, DFT, and PCA. The performance of the CSVM and SIMCA classifiers within the ECG classification context is discussed. An evaluation of the performance of each classification scheme is also provided. Finally, this chapter provides a discussion of the overall classification framework that was used across all studies. A discussion of the results in terms of the chosen classification performance metrics is provided; this is compared with other studies of ECG classifiers found in the literature. Some of the results in this chapter were included in the candidate's publications [28-29].

*Chapter 6* provides a summary of the project's outcomes and some concluding remarked. Moreover, recommendations for possible further developments and future work are mentioned. The possibility of using mobile phone technologies for analysing ECG signals is also discussed. Furthermore, it discusses the use of Clifford algebra in data fusion as well as the potential use of Clifford algebra support vector machines and geometric neurons for multi-lead analysis and ECG beat classification in the future.

# **Chapter 2     ECG Signal Characterisation Studies and Identification of Arrhythmias in the ECG Signal**

## **2.1 Introduction**

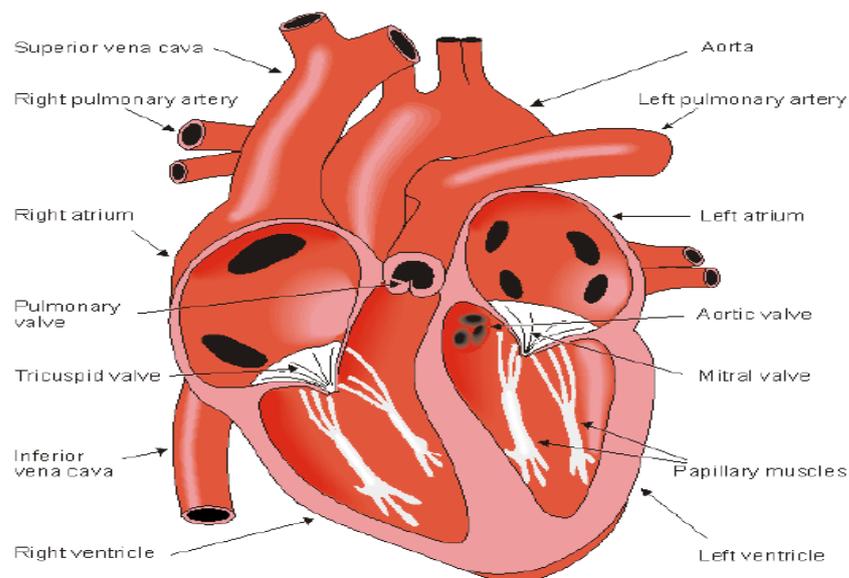
An ECG is a graphic representation of the heart muscle's electrical activity. The contraction of any heart muscle is associated with electrical changes called depolarisation and these changes can be detected by electrodes attached to the surface of the body. The ECG monitor provides the necessary visualisation [30]. In all ECG records, there are 12 ECG data acquisition channels that must be considered, and information from all of these may be used in the process of determining cardiac diseases. The ECG record provides an insight into pathological diseases of the heart based on the heart's electrical activity. From a physician's perspective, it is also imperative to understand the orientation of the planes associated with the propagating waves, which are associated with each of the leads' recordings. In this way, the activity of the heart can be mapped. Recording strips also provide details of the activity of the heart between the positive and negative pole. The ECG record provides evidence and information that can help a doctor to diagnose abnormalities in the heart rate such as arrhythmias, myocardial infarctions, atrial enlargements, ventricular hypertrophies, and bundle branch blocks. A patient's ECG record can be used to detect abnormal cardiac rhythm, abnormal cardiac conduction, ischaemia of the myocardium, and hypertrophy, and it can be used to provide a dynamic picture of how the heart is functioning as a pump. Heart function tests are necessary for an accurate assessment of heart disease risk. In the following section of this chapter, heart function and the interpretation of the ECG are discussed further.

## **2.2 The anatomy of the heart and its function**

### **2.2.1 The heart's anatomy**

The main function of the cardiovascular system is to transport nutrients and oxygen to the entire body. The heart is a muscular organ that pumps blood to all the tissues in the body through a network of blood vessels [31]. The network can be divided into two circuits: the pulmonary circuit and the systemic circuit. The pulmonary circuit carries carbon dioxide-rich blood from the heart to the surfaces of the lungs and returns oxygen-rich blood to the heart.

The systemic circuit transports oxygen-rich blood from the heart to the rest of the body's cells, returning carbon dioxide-rich blood back to the heart. The heart has four chambers and several atrioventricular and sinoatrial nodes in the atrium, as shown in Fig. 2.1. The upper two chambers are called the right atrium and left atrium, whereas the lower two are called the right ventricle and left ventricle. These chambers are separated by a wall of tissue called the septum, which separates the left from the right side of the heart. The right atrium receives blood from the systemic circuit while the right ventricle discharges blood into the pulmonary circuit [32]. The right side of the heart pumps blood through the lungs where it takes in oxygen, while the left side of the heart receives the blood containing oxygen and pumps the blood to the rest of the body. The right atrium receives oxygen-devoid blood from the body and the left atrium receives oxygen-rich blood from the lungs. In other words, the purpose of the atria is to receive blood from the body. The atria are separated from the ventricles by the tricuspid valve on the right side and the mitral valve on the left side. The blood flows from the atria into the ventricles when these valves are opened. The ventricles are made of stronger muscle tissue than the atria because their function is to pump blood throughout the body. The right ventricle pumps the oxygen-devoid blood to the lungs to absorb oxygen and release carbon dioxide, while the left ventricle pumps the oxygen-rich blood to the body's organs [33]. The atria are attached to the ventricles by fibrous, non-conductive tissue, which keeps the ventricles electrically isolated from the atria.



**Figure 2.1: Anatomy of the human heart, with chambers and valves, adopted from [34]**

The human body has about 5.6 litres (6 quarts) of blood, all of which circulate through the body three times every minute. The right atrium and the right ventricle together form a pump to circulate blood to the lungs. The heart's cycle begins when oxygen-poor blood from the body moves into the right atrium. Oxygen-poor blood is received through large veins called the superior and inferior vena cava and flows into the right atrium. Then the blood flows through the right atrium into the right ventricle that serves as a pump, which sends the blood to the lungs. The blood releases waste gases and picks up oxygen within the lungs. This newly oxygen-rich blood returns from the lungs to the left atrium through the pulmonary circuit. Then the blood flows through the left atrium into the left ventricle. Finally, the left ventricle pumps the oxygen-rich blood out through the aorta, and from there to all parts of the body [31].

### **2.2.2 Heart dynamics as related to ECG recordings**

The mechanical pumping action of the heart results from electrical activation that is applied to the spread of electrical signals through the atria and ventricles as shown in Fig 2.2. At initiation, the signal for the heartbeat begins in the sinus or sinoatrial (SA) node, which is located in the right atrium near the opening of the superior vena cava. The SA node is a small collection of specialised cells capable of automatically generating an electrical stimulus. This stimulus spreads first through the right atrium from the SA node and then into the left atrium. Next, the electrical stimulus reaches specialised conduction tissues in the atrioventricular (AV) junction that act as an electrical relay connecting the atria and ventricles. The AV junction is located at the base of the interatrial septum and extends into the interventricular septum. The upper part of the AV junction is the AV node, while the lower part of the AV junction is called the bundle of His. The bundle of His divides into two main branches: the right bundle branch and left bundle branch. The right bundle branch distributes the stimulus to the right ventricle, whereas the left bundle branch distributes the stimulus to the left ventricle. Finally, the electrical signal then spreads simultaneously down the left and right bundle branches into the ventricular myocardium (ventricular muscle) through specialised conducting cells called Purkinje fibres [1].

The ECG is capable of recording only relatively large currents produced by the mass of working (pumping) heart muscle. The ECG records the average electrical activity that corresponds to a large mass of atrial and ventricular cells. Each cardiac electrical activity

phase is identified by a specific wave labelled alphabetically P, QRS and T. In order to record the electrical activity of the heart, 12 standard ECG leads are placed on the surface of the body. These leads detect the changes in voltage of the cardiac muscle.

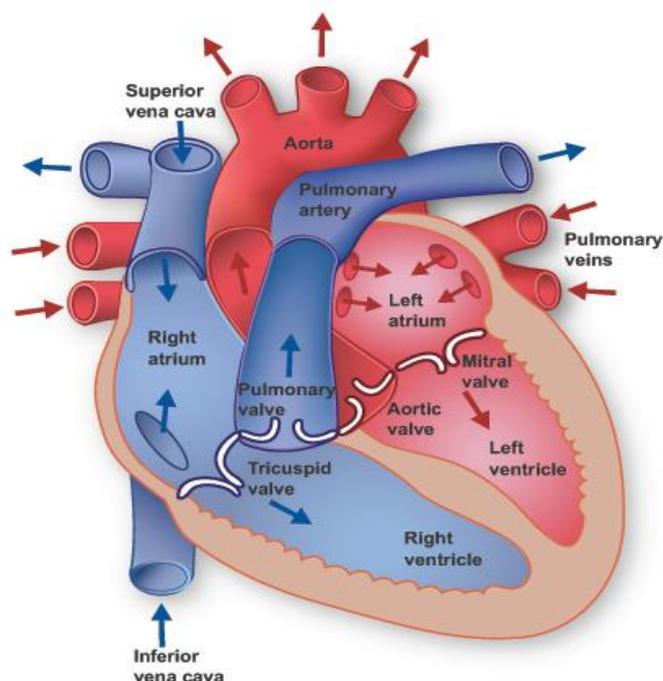


Figure 2.2: The mechanical pumping action of the heart, adopted from [35]

### 2.3 A concise historical account of work leading to today's recording practice and interpretation of the ECG signals

The first ECG recording was made by Willem Einthoven in 1895. In his work, the P, QRS, T waves were also defined for the first time [36]. The 1930s saw the invention of the unipolar extremity leads and the six precordial leads. These were added to the original three bipolar extremity leads by Dr Frank N. Wilson and his colleagues at the University of Michigan. Originally, three augmented unipolar extremity leads were used: aVR, aVL, and aVF. These were invented by Dr Emanuel Goldberger. Currently, 12 leads are normally employed, including the six limb leads (I, II, III, aVR, aVL and aVF) and the six precordial leads (V1 to V6). In addition to this, the basic ECG waves are separately labelled as the P wave, the QRS complex, the T wave, and the U wave. In this way, the analysis of ECG waves is used to obtain a dynamic picture of the heart's function and arrhythmias.

### **2.3.1 The standard 12-lead ECG record**

The multi-lead ECG is a medical device, which records the heart's electrical activity with electrodes placed on the skin in specific locations. It consists of the measurement of electrical activity on the body surface associated with myocardial contraction with respect to time. The standard ECG contains 12 different leads which record the same electric events but from a different view. The 12 leads provide a three-dimensional view of the electrical activity of the heart. The determination of the proper placement of these leads is very important, as this dictates the accuracy of the reading in the ECG grids. In fact, one should always consider that monitoring the cardiac output involves a series of cardiac cycles, as records associated with a single cycle are insufficient to interpret problems with the heart's function. In other words, the multiple leads are very important when it comes to an ECG because they provide 12 different views of cardiac activity, by showing the direction of wave deflection as related to the heart's electrical activity [37]. The standard ECG recording electrodes are placed on the arms, legs, and precordial wall. ECG patterns are obtained when electrodes are placed at various points on the chest. These leads can be divided into two groups of six on limbs (extremity) and the chest (precordial) [36]. The limb leads are used to record voltages on the frontal plane of the body, while the six precordial leads record voltages on the horizontal plane.

#### **2.3.1.1 ECG records using limb leads**

The six limb leads record voltage differences by means of electrodes placed on the extremities. According to [1], the limb leads can be divided into two subgroups depending on their historical development; three leads on each standard bipolar limb (I, II, and III) and three augmented unipolar limb leads (aVR, aVL, and aVF), as shown in Fig. 2.3. The purpose of bipolar limb leads is to calculate the mean depolarisation vector of the heart in the frontal plane. The unipolar augmented limb leads are used to determine the orientation of the heart [38]. The limb leads are connected to metal electrodes, which are placed on the arms and legs of the patient. In addition, the arm electrodes need to be attached just above the wrist and the leg electrodes need to be attached above the ankles, as shown in Fig. 2.3. Lead I, II and VL look at the left lateral surface of the heart, while lead III and VF look at the inferior surface and lead VR looks at the right atrium.

Lead I records the difference in voltage between the left arm (LA) and right arm (RA) electrodes, Lead II records the difference between the left leg (LL) and right arm (RA)

electrodes, whilst Lead III records the difference between the left leg (LL) and left arm (LA) electrodes. These leads are used to provide a basic ECG recording [36].

Limb leads are very important as they allow physicians to assess whether there is an ischaemic event occurring in the frontal area of the heart. These leads also represent a way of detecting if there is an electrical impulse in this area of the heart. In this case, determining the current flow in the frontal area of the heart indicates whether or not the heart is in good condition. This in turn indicates that there is nothing happening in the heart. It is an indication that there is no myocardial infarction or heart blockage occurring. Indeed, electrical impulse detection is confirmatory evidence that a certain cardiac disease is present in the heart, and more specifically in the frontal or on the inner area of the heart. In this case, proper and precise reading in this grid is very important for the accurate diagnosis of the patient's heart condition [39].

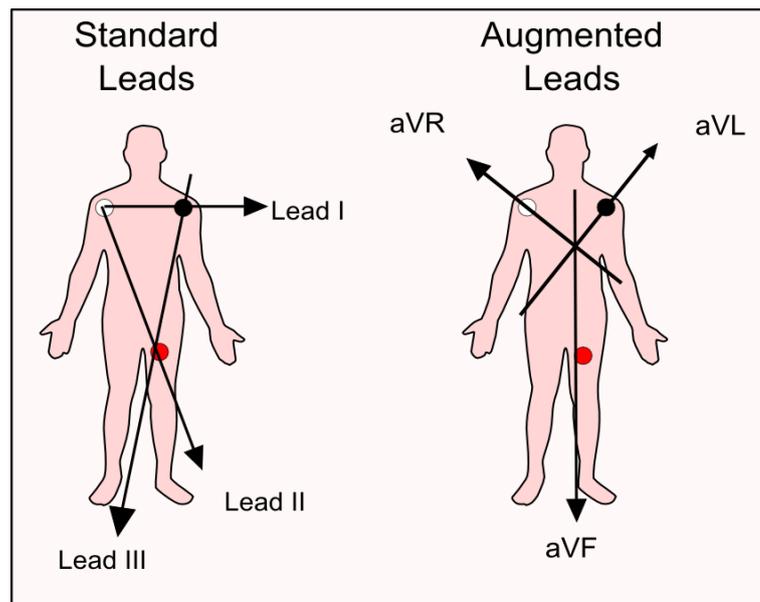


Figure 2.3: Standard and augmented limb leads positions, adopted from [40]

### 2.3.1.2 ECG records using the precordial leads

The precordial leads record voltage differences by means of electrodes placed at several positions on the chest wall, as illustrated in Fig. 2.4. These are actually precordial leads which measure the propagation of voltages across the horizontal plane of the heart [41].

Lead V1 is located in the fourth right intercostal space to the right of the sternum; lead V2 is located in the fourth left intercostal space to the left of the sternum. Both V1 and V2

leads record voltage from the right ventricle. V3 and V4 leads record voltage from the anterior surface and lateral walls of the left ventricle, while V5 and V6 record voltage from the frontal and lateral walls of the left ventricle. Lead V3 is placed on a line central to leads V2 and V4, whereas lead V4 is placed in the mid-clavicular line in the fifth interspace. Lead V5 is placed in the anterior axillary line at the same level as lead V4, and lead V6 is placed in the mid axillary line at the same level as lead V4 [36]. Consequently, the six limb leads record electrical voltages transmitted to the frontal or vertical plane of the body, whereas the precordial leads record voltages transmitted onto the horizontal plane from the front and the left side. A total of 12 leads offer three-dimensional visualisations of atrial and ventricular depolarisation and repolarisation.

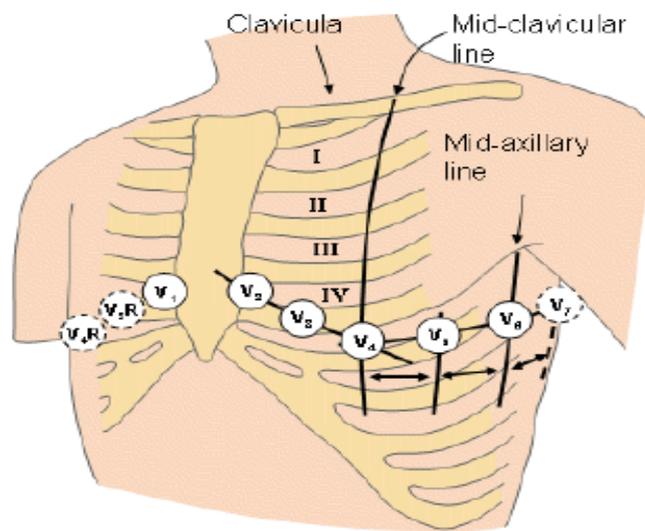
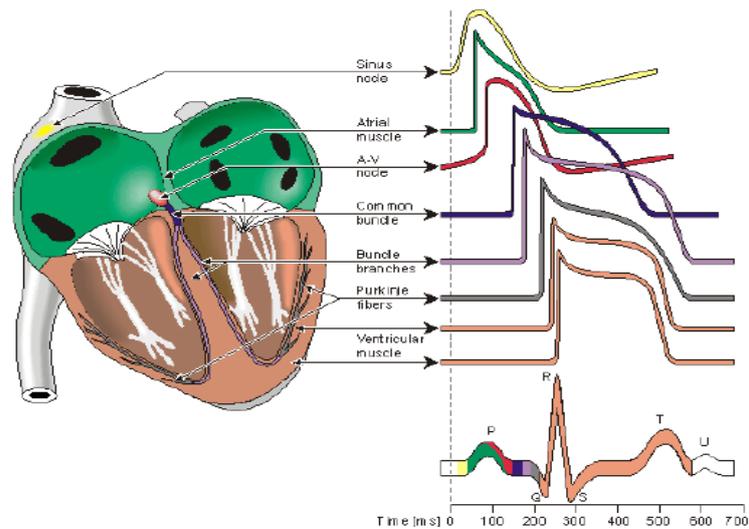


Figure 2.4: Precordial leads (electrodes) positions, adopted from [34]

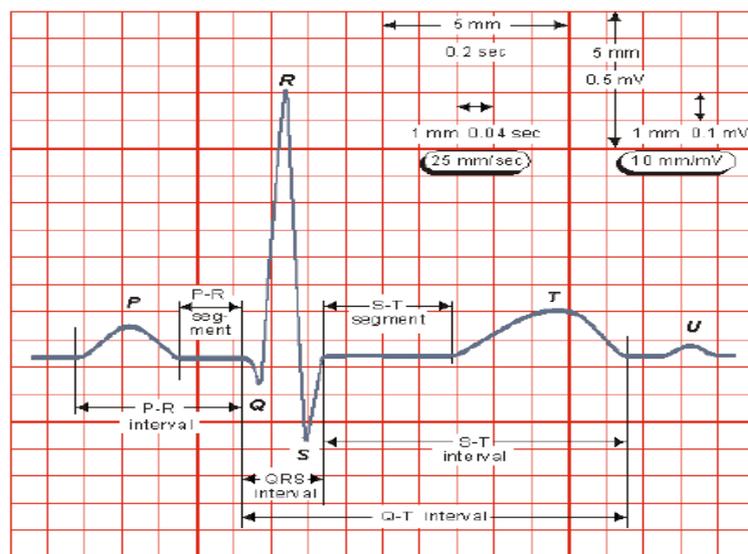
### 2.3.2 Description of the various ECG wave types recorded

The cardiac activation process of the heart muscle is performed through two different processes: depolarisation (the spread of a stimulus through the heart muscle) and repolarisation (where heart muscle cells return to their resting state). The cardiac cycle initiates with the P wave, which corresponds to the period of atrial depolarisation in the heart. The QRS complex is the second wave, and it corresponds to the period of ventricular depolarisation. The start and end points of the QRS complex are denoted as the Q and J points, respectively. The most noticeable feature of an ECG waveform is the QRS complex. The T wave presents the period of ventricular repolarisation, and the end point of this wave represents the end of the cardiac cycle as shown in Fig. 2.5 [42].



**Figure 2.5: The origin of the electrical activity and electrophysiology of the heart, adopted from [34]**

Fig. 2.6 shows the basic characteristic features of the ECG wave, showing the interval and associated segment in a single beat record. Some intervals and segments between waves are identified as QT, ST and PR features. The P-QRS-T sequence is usually recorded on special ECG graph paper, which provides a grid for visual identification of the duration of the features. The ECG grid resolution is 1 mm squares. Most often, this corresponds to time duration of 0.04 s. The ECG vertical axis relates to the voltages [39].



**Figure 2.6: The basic ECG wave, interval, and segment, adopted from [34]**

### **2.3.2.1 Physiological relevance and good practice in recording the P wave**

The P wave corresponds to the atrial depolarisation that occurs inside the heart whenever there is a new heartbeat. In this process, it is the first waveform in the ECG following a QRS complex. The normal characteristics of this are an amplitude of 2 to 3 mm high, a duration of 0.06 to 0.12 s and a shape that is round and upright. The frontal plane of the P wave axis should be between  $0^\circ$  and  $+75^\circ$  and the voltage in the limb leads should not exceed 0.25 and 0.15 mV in the precordial leads. Usually, in most leads, the P wave is positive, whereas it is always negative in one of the limb leads (aVR) and sometimes can have a negative deflection in leads III and V1 [38].

### **2.3.2.2 Typical characteristics of the QRS complex**

There is a large deflection associated with the ECG signal when the ventricles are depolarised. This is known as the QRS complex wave. This contains more than two or three deflections. It is the ventricular depolarisation which would occur after this for every single contraction which takes place in the heart. The muscle mass of the atria is small compared with that of the ventricles, and as a consequence the P wave is smaller than the QRS complex. The normal duration does not exceed 0.12 s and has a typical duration of 0.06 to 0.10 s. The typical voltage usually varies between 1.5 and 2.0 mV. If the duration of the QRS complex is greater than 0.12 s, it is due to an asynchronous depolarisation of both ventricles. The QRS complex follows the PR interval and has an amplitude of 5 to 30 mm high, although this may differ depending on the lead used [38], as voltage gain settings in the transducer can vary.

### **2.3.2.3 Typical characteristics of the T wave and its physiological relevance**

A T wave is associated with the return of the ventricular mass to its resting electrical state; a process known as repolarisation. In other words, the T wave represents ventricular recovery. This has an amplitude of 0.5 to 10 mm and takes place at every heartbeat. The shape of the T wave is round and smooth, usually upright, and maybe inverted depending on the lead's polarity. This particular wave follows the S wave of the heart, which can be determined by the attending cardiologist [43].

#### **2.3.2.4 Typical characteristics of the U wave and its physiological relevance**

A U wave corresponds to a small deflection, which is sometimes seen after the T wave and represents the last phase of ventricular repolarisation. This is the recovery period of the Purkinje fibres or the ventricular conduction fibres. This wave follows the T wave and has a shape that is typically rounded and upright. The direction of the U wave is most often the same as that of the T wave, but sometimes a negative U wave can appear with positive T waves. This wave can be identified in about one-quarter of the population, but might not appear on the recording strip sufficiently clearly. It does not provide any significant information with regard to the activity of the heart [43].

#### **2.3.2.5 Typical characteristics of the J wave and its physiological relevance**

A J wave refers to the elevated point of an electrocardiogram, which appears at the very beginning of the ST segment. It is also referred to as the Osborn wave. This wave appears as a late delta wave, which usually follows the QRS complex. Some doctors feel that this wave is in fact a smaller secondary wave. Whenever there is an evident case of an identified J wave period, it can be regarded as a pathognomonic sign of hypothermia or hyperkalaemia [43].

### **2.3.3 The explanation of the physiological relevance of ECG segments and intervals**

Segments and intervals are other identifying elements in the ECG waveform. ECG segments relate to the duration of the isoelectric line between waves while ECG intervals correspond to the time elapsed between the same segments of adjacent waves. A further aim of an ECG analysis is to evaluate the shape, amplitude, and duration of the waves, segments, and intervals [38].

ECG intervals measure the depolarisation and repolarisation duration, which can be observed in the heart for every activity and provide an assessment of overall cardiovascular integrity, which occur in every contraction. When the heart beats faster, one can easily notice that the length of the interval is shorter; when the heart beats slower, the duration is longer. The most important feature to consider in the ECG is the QT interval. If there is a change in the length of each heart interval, it would indicate that there is something wrong with the physiological integrity of the heart. Identification of this particular abnormality provides an

additional assessment of overall heart health and so could lead to a new way of making medical assessments.

### **2.3.3.1 Physiological relevance of the PR interval**

The PR interval is defined as the dead time between the end of the P wave and the beginning of the QRS complex. This tracks the propagation of an electrical impulse from the atria to the AV node across all pathways (P), to the end point, which is the bundle branches. The duration of this particular wave is 0.12 to 0.20 s [38].

### **2.3.3.2 Physiological relevance of the ST segment**

The observed segments in an ECG signal depict the end of a contraction and the beginning of the following one. They confirm the continuity of the electrical activity manifested in the heart. More specifically, they denote the end of atrial and ventricular contraction and also the relaxation of the atria and ventricles in preparation of another contraction to be initiated. The feature associated with ECG rhythmic strip is the ST segments that represent the end of ventricular contraction and depolarisation and the beginning of ventricular recovery.

The ST segment is particularly important to a physiologist, as it reveals all related physiological responses that are essential to interpreting the general degree of health of the patient's heart [44]. The characteristics of the ST segment are that it extends from the end of the S wave to the beginning of the T wave. The deflection is isoelectric, with a duration of 0.5 to 1 mm for every heartbeat [43].

### **2.3.3.3 Physiological relevance of the QT interval**

The QT interval is associated with the return of stimulated ventricles to their resting state known as ventricular repolarisation. In addition, the QT interval provides an estimate of the ventricular depolarisation and repolarisation as revealed by the electrocardiogram. This interval occurs at the beginning of the QRS complex and extends to the end of the T wave. Normally, it has duration of between 0.36 and 0.44 s [12].

### **2.3.3.4 Physiological relevance of the PQ interval and ST-T interval**

The PQ interval shows the time elapsed between atrial depolarisation and the onset of ventricular depolarisation. Its normal range is between 0.12 and 0.20 s. The ST-T interval

provides a measurement of the speed of the repolarisation of the ventricular muscle. It is a part of the QRS complex (the J point) and lasts until the onset of the T wave [38].

## **2.4 Heart rhythms and detection of arrhythmias in the ECG signal**

Arrhythmia is a cardiac condition caused by the abnormal electrical activity of the heart. According to [45], arrhythmias can be divided into two groups. The first group relates to life-threatening arrhythmias that require immediate therapy with a defibrillator. Arrhythmias are the result of ventricular fibrillation and tachycardia. Many solutions with high sensitivity and specificity have been developed for detection of these arrhythmias. The second group of arrhythmias are not imminently life threatening but may require therapy to prevent additional problems. This study focuses on analysing and detecting arrhythmias from this second group such as premature ventricular contraction (PVC) and fusion of ventricular and normal beats.

The normal rhythm of the heart where there is no disease or disorder in the morphology of the ECG signal is called the normal sinus rhythm (NSR). The heart rate is generally characterized by regular beats at a frequency of 60 to 100 beats per minute. The rhythm is known as sinus tachycardia when the heart rate is over 100 beats per minute. However, the rhythm is called bradycardia when the heart rate is too slow [46].

Typical heart rates differ with every person, for example because of cardiovascular conditioning, so correlations of heartbeat rates can be misleading. Unusual heart rates have variable causes and affect individuals in various ways. Heartbeat issues, called arrhythmias, are one paradox of medication. Just about anybody's heart can produce an extra beat or two, and the upsetting manifestations that might go with the additional pulsations like palpitations do not often cause a serious issue. Yet an undetected arrhythmia, likewise, might set off a chain of events prompting sudden death from heart failure. In the following section, discussion of some types of arrhythmias is provided.

### **2.4.1 Sinus node arrhythmias**

This sort of arrhythmia is associated with the SA node of the heart. A typical feature of these arrhythmias is that the P wave morphology of the ECG is normal. Sinus arrhythmia, sinus bradycardia, and sinus arrest are some manifestations of these arrhythmias [46].

## 2.4.2 Myocardial infarction and ischaemia

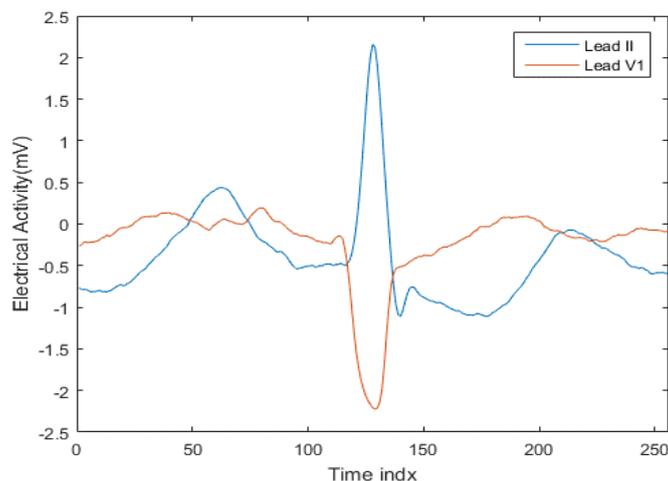
Myocardial dead tissue occurs following myocardial ischaemia. It is a result of a reduced blood supply to the heart, below a certain threshold that overpowers myocardial cellular repair systems intended to keep up typical working function and homeostasis. Ischaemia often results in irreversible myocardial cell damage or death. An interruption in the supply of myocardial oxygen happens when a thrombus is superimposed on ulcerated or shaky atherosclerotic plaque. This often results in coronary impediment [47].

## 2.4.3 Atrial arrhythmias

Irregular rhythms are generated by the left or right atrium in the human heart. The main cause of this arrhythmia is often old age. The most predominant changes in the P wave and Q wave features relate to the atrial electrical activities [48]. Atrial arrhythmias start outside the SA node but inside the atria. There are different arrhythmia types such as premature atrial contractions (PAC) and atrial flutter and atrial fibrillation.

During PAC arrhythmias, an irregular P wave morphology is produced after a normal QRS complex and T wave. This happens because an ectopic pacemaker terminates before the SA node. PVC has a signature morphological feature: before the QRS complex, an extra beat will be present, as shown in Fig. 2.7. In addition, the P-R segment will be non-isoelectric as a result of the extra beat [48]. PACs might happen as a doublet where two PACs are produced continuously. Whenever three or more successive PACs happen, the patient is thought to be suffering from atrial tachycardia. In atrial tachycardia, the heart rate is quick and ranges from 160 to 240 beats per minute.

During atrial flutter (AF), the atrial beat rate is very high, increasing from 240 to 360 beats per minute. The irregular P waves happen frequently and thus they rapidly take the morphology of a saw-tooth waveform, which is called ripple (F) waves. AF happens when a single group of atrial muscles triggers an electrical pulse that disrupts the regular impulse from the SA node. The features of the waveform are regular R-R intervals and there are saw-tooth waveforms present between R peaks [48]. The AF rate is even faster than in atrial tachycardia, which is between 160 and 240 beats per minute.



**Figure 2.7: The PAC beat were taken from the MIT-BIH Arrhythmia database**

During AF, the atrial rate sometimes surpasses 350 beats per minute. This arrhythmia happens as a result of compression of various parts of the atria. The higher atrial beat rate and compression prompts insufficient pumping of blood into the ventricles [46]. In other words, atrial fibrillation is initiated by multiple groups of the atrial muscle, triggering pulses at random, causing the sinus node's regular electrical impulse to be disrupted, and thus it fails to trigger the AV node regularly. The main features of AF are the irregular R-R intervals; furthermore, both P wave and T wave are unrecognizable [48].

#### 2.4.4 Junctional arrhythmias

A junctional rhythm happens when the primary pacemaker, the SA node, fails to send the electrical impulses along the presented pathway. Therefore, the secondary pacemaker, the AV node, becomes the main pacemaker of the heart. The features of this rhythm are a regular R-R interval and a delay in the P wave that appears after the R waves [48].

Junctional ectopic tachycardia is an irregular programmed tachyarrhythmia that emerges in the atrioventricular conduction system over the bifurcation of the bundle of His. As a result of this arrhythmia, the P wave morphology is abnormal. The abnormal P wave would be opposite the P wave of the normal sinus due to depolarization. This wave circulates in the opposite direction from the AV node towards the atria [46]. Children can have junctional ectopic tachycardia as an inherent sickness or other physiological imperfections, as

a result of cardiac surgery. This arrhythmia is uncommon in grown-ups. Numerous youngsters with inborn junctional ectopic tachycardia have a family history of this type of arrhythmia.

Innate junctional ectopic tachycardia for the most part occurs at the same time in life. At the point when the arrhythmia follows cardiac surgery, the youngsters regularly have operations including atrioventricular intersections. Congestive heart failure and hypotension involve fast heart rates.

The trademark electrocardiographic signature is atrioventricular separation. Inherent junctional ectopic tachycardia is typically treated with beta-adrenergic blocking drugs and amiodarone, which will diminish the ventricular rate. Atrial pacing can briefly re-establish atrioventricular synchrony and increase cardiac yield. Junctional ectopic tachycardia can significantly compromise cardiac function when it is the result of cardiac surgery, as a rule in kids with inherent heart sickness. Amiodarone, beta-adrenergic blocking drugs and propafenone diminish the rate of the tachycardia; however, other antiarrhythmic medications are generally insufficient. Removal of the atrioventricular junctional and ventricular pacing might sometimes be required. This arrhythmia can be a reliable proxy to determine, most of the time, whether the patient will survive the early post-operative period [5].

### 2.4.5 Ventricular arrhythmias

The general characteristic of ventricular arrhythmias is that the QRS-complex is wider, as shown in Fig 2.8. Ventricular arrhythmias can be associated with the following conditions: untimely ventricular contractions (PVC), ventricular tachycardia (VT) and ventricular fibrillation (VF).

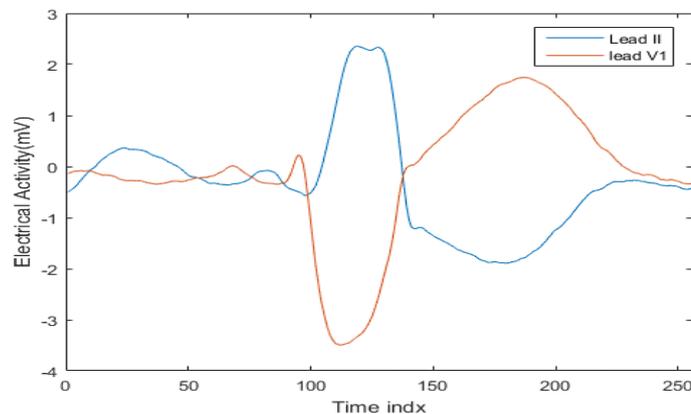


Figure 2.8: The PVC beats were taken from the MIT-BIH Arrhythmia database

In PVC arrhythmias there is no depolarisation of the atria or the SA hub. The morphology of P waves maintains their underlying rhythm. This kind of ventricular arrhythmia might happen at any place in the heartbeat cycle. PVCs manifest as segregates on the off events that occur separately or as couplets if two continuous PVCs occur.

During ventricular tachycardia, the heart rate is between 110 and 250 beats per minute. The QRS complex in VT is also strangely wide. This arrhythmia is considered life-debilitating as the fast rate might lead to ventricular filling, resulting in a drop in cardiac yield. The common features in VT are the wide R waves and short R-R intervals.

Ventricular fibrillation happens when various ectopic pacemakers in the ventricles cause diverse parts of the myocardium to contract at various times in a non-synchronised manner. The ventricular vacillation shows an exceptionally quick ventricular rate with a saw-tooth-like ECG waveform [46]. The features of this arrhythmia are noise and a quick heart rate of 300 beats per minute (sometimes this can even reach more than 300 beats per minute).

#### **2.4.6 Atrioventricular blocks**

Another group of arrhythmias that can be detected from the ECG signal is heart blocks. Atrioventricular block refers to the typical pathological spread of the electrical activity along the conduction pathways to the ventricles. Such a blockage might delay the propagation of the waves or totally incapacitate whatever is left of the conduction system. There are three types of blockage: first-degree block, second-degree block and third-degree block. A first-degree AV block happens when all the P waves are led to the ventricles; however, the PR interval is delayed. The delay is revealed as a feature on the ECG beat analysis that shows an extended P-R interval. Second-degree AV blocks happen when a portion of the P waves fails to reach the ventricles. The main feature of this block is an irregular R-R interval. Furthermore, the QRS complexes in the ECG record are missing consecutively. During third-degree AV blocks, the tempo of the P waves is totally separated from the QRS event. Every beat occurs at its own particular rate.

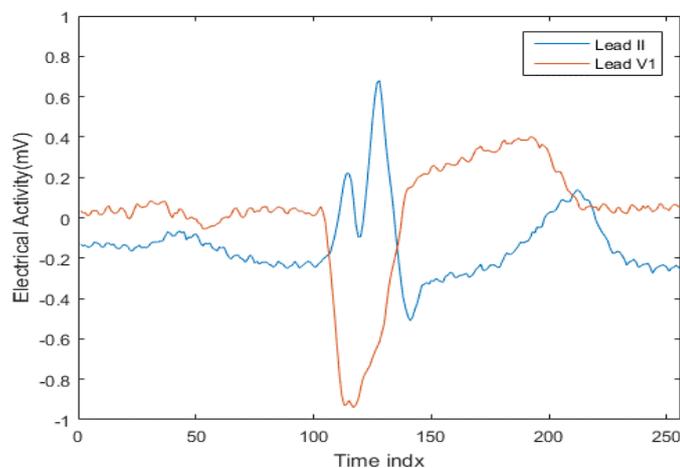
The bundle branch block stops the conduction of the electrical stimulation from the AV node to the entire conduction system. Because of this blockage, there might be localised myocardial necrosis needing cardiac surgery. The bundle branch block beat is classed as being of two types. These are the left bundle branch block beat (LBBB) and right bundle

branch block beat (RBBB). In LBBB, the left bundle branch will keep the electrical stimulation forces from the AV hub from depolarizing the left ventricular myocardium in the ordinary way. At the point when the right bundle branch is blocked, the electrical excitation from the AV hub is not capable depolarizing the ventricular myocardium [46].

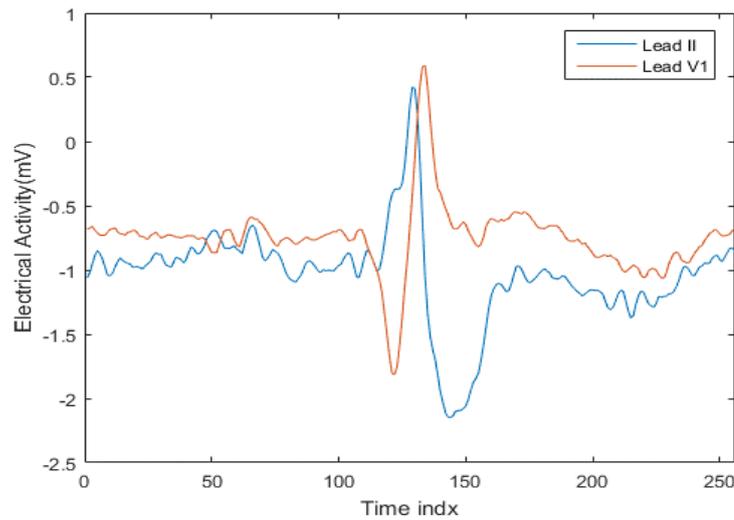
### 2.4.7 Bundle branch blocks (BBB)

The ventricles of the heart (either left or right) contain an adequate muscle cell mass; viable depolarization of the considerable number of cells requires a specific conduction pathway inside of the ventricle. The bundle branches (left and right) correspond morphologically to the first division in the ventricular conduction system after the bundle of His. Conduction blocks can happen in either of the two bundle branches. As noted, these can happen as the after-effect of localized necrosis of the tissue, in spite of the fact that various generally typical individuals may have a bundle branch block because of the fill-up of the conduction pathway with sinewy tissue [49].

The LBBB is normally associated with poor intraventricular conduction. The common surface ECG highlight of LBBB is a prolongation of the QRS complex event above 0.11 s in conjunction with a delay of the natural redirection in leads V5 and V6 of more than 60 ms and no septal Q waves in leads I, V5, and V6 because of the unusual septal actuation from right to left. LBBB might instigate irregularities in left ventricular function because of anomalous off-beat constriction patterns which can be fixed by bi-ventricular pacing (resynchronization treatment) [50]. Fig 2.9 and 2.10 show the LBBB and RBBB beats were taken from the MIT-BIH Arrhythmia database lead II and lead V1 respectively.



**Figure 2.9: The LBBB beats were taken from the MIT-BIH Arrhythmia database**



**Figure 2.10: The RBBB beats were taken from the MIT-BIH Arrhythmia database**

## 2.5 Description of the ECG databases used in the current study

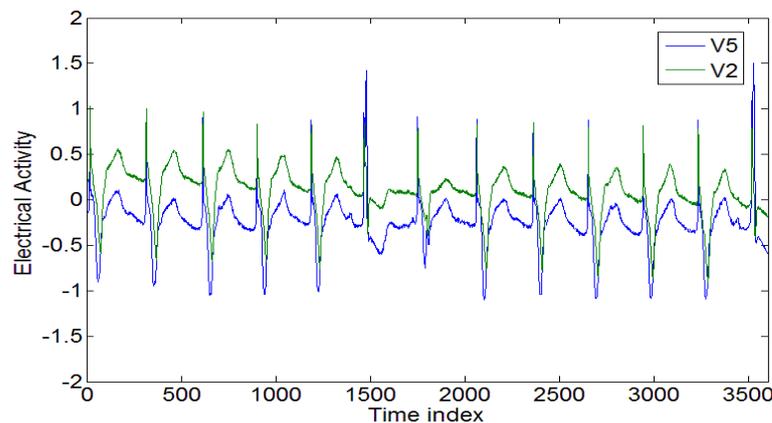
### 2.5.1 The MIT-BIH Arrhythmia Database

The MIT/BIH Arrhythmia Database is used in the majority of studies for software-based performance evaluations. This database was the first to be established and provides standard test records for assessment of arrhythmia identifiers. It has been utilized to assess cardiac flow at more than 500 labs around the world. Initially, the database was stored on a nine-track computerised tape at 800 and 1,600 bpi, and on quarter-inch IRIG-position FM simple tape. In August 1989, a CD-ROM rendition of the database was created.

The MIT-BIH Arrhythmia Database contains 48 half-hour portions of two-channel wandering ECG recordings, acquired from 47 subjects. Recordings were made between 1975 and 1979. Each of these recordings includes two-lead information such as from limb lead II and one from the precordial leads V1, V2, V4 or V5 or two precordial leads as presented in Fig. 2.11. Twenty-three recordings were picked randomly from 24-hour ambulatory ECG recordings gathered from a blended populace of inpatients (around 60%) and outpatients (around 40%) at Boston's Beth Israel Hospital.

In addition, the recordings were digitized at 360 samples for every second per channel with 11-bit resolution over a 10 mV range. Two or more cardiologists freely clarified every record; disagreements were resolved to obtain the computer-readable reference annotations for every beat (around 110,000 annotations) included within the database [51],[27].

There are 12 types of ECG waveforms of abnormal beats in this database: left bundle branch block (LBBB), right bundle branch block (RBBB), atrial premature beat (APC), aberrated atrial premature beat (a), nodal (junctional) premature beat (J), ventricular premature beat (VPC), fusion of ventricular and normal beat (FUSION), ventricular flutter wave (I), nodal (junctional) escape beat (j), ventricular escape beat (E), supraventricular premature beat (S), and fusion of paced and normal beat (PFUS), and the waveforms corresponding to the normal sinus rhythm (N). The vertical axis  $y$  is measured in  $\mu\text{V}$  and presents amplitude of ECG waveforms, while the horizontal axis  $x$  scaled as the number of points (at a 360 Hz sampling rate; 360 samples corresponds to every second and one point corresponds to approximately 2.8 ms), shows time index (duration) of ECG waveforms [44]. The diagram in Fig. 2.11 illustrates ECG waveforms and beat that is taken from the MIT-BIH Arrhythmia Database patient record number 104.

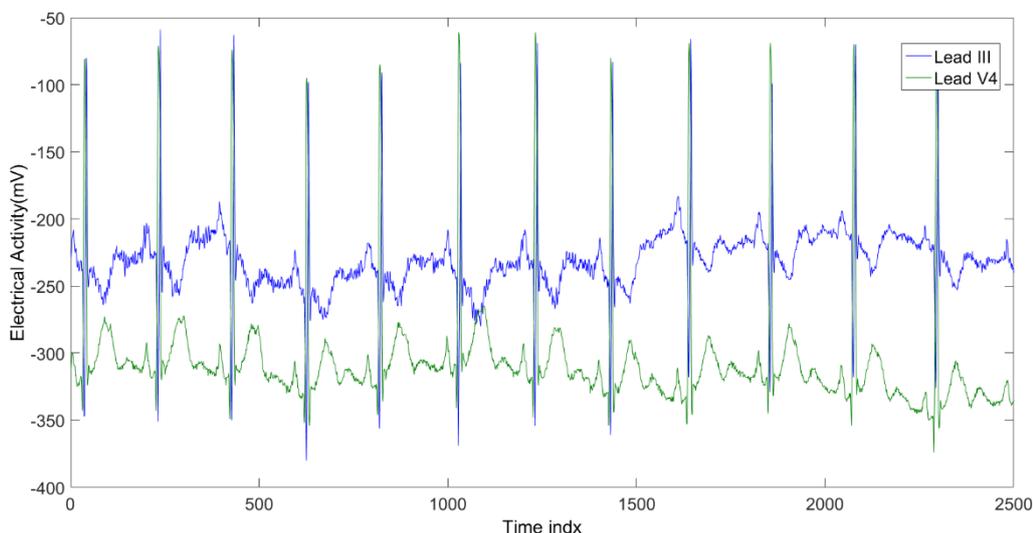


**Figure 2.11:** A segment of an ECG signal taken from the MIT-BIH Arrhythmia database (record no 104)

## 2.5.2 The European ST-T Database

The European ST-T Database is often utilised in order to investigate ST and T wave changes. This database comprises 90 annotated selections of ambulatory ECG recordings from 79 subjects; data was lost for one subject. The subjects included 70 mature men aged 30 to 84, and 8 mature ladies aged 55 to 71. Records e0118–e0122 originated from the same subject as did records e0123–e0126, records e0129 and e0133, records e0136 and e0139, records e0147 and e0148, records e0154 and e0155, and records e0162 and e0163. Then myocardial ischaemia was diagnosed or suspected for every subject; this was determined using the following procedure: firstly, multiple selection criteria were set up in order to obtain a representative selection of ECG ischaemia events, including baseline ST segment

displacement resulting from hypertension and ventricular dyskinesia. Moreover, this database incorporated 367 episodes of ST segment change, 401 episodes of T wave change with periods spanning from 30 s to a few minutes, and crest relocations extending from 100  $\mu\text{V}$  to more than one mV. In addition to this, 11 episodes of axis movement bringing about evident ST change and 10 episodes of axis movement bringing about clear T wave change were stamped. Minimal clinical reports were used to archive every record. These reports are contained inside the header (.hea) documents and are connected with every record. They condense pathology, pharmaceuticals, electrolyte imbalance, and technical data for every recording [27]. The diagram in Fig. 2.12 illustrates ECG waveforms and beat that is taken from the European ST-T database record no e0104.



**Figure 2.12: A segment of an ECG signal taken from the European ST-T database (record no e0104) depicting lead III and V4 information**

### 2.5.3 The St Petersburg INCART 12-lead arrhythmia database (incartdb)

This database comprises 75 annotated recordings removed from 32 Holter records. Each one of these records is 30 minutes in length and contains 12 standard leads, each sampled at 257 Hz, with increases shifting from 250 to 1,100 simple to-advanced converter units per millivolt. Pick-ups for every record are indicated in the document. The annotated reference documents contained more than 175,000 beat annotations in total. The first records were gathered from patients undergoing tests for coronary vein illness (17 men and 15 women, aged 18–80; mean age: 58). In addition, none of the patients had pacemakers; most had

ventricular ectopic beats. In selecting records to be incorporated into the database, preference was given to subjects with ECGs predicting ischemia, coronary vein illness, conduction variations from the norm, and arrhythmias [27].

## **2.6 Summary**

In this chapter, the basic features of ECG recording and their physiological relevance were discussed. It was explained how such recordings can diagnose a patient with cardiac ischaemia and infarction. Interpretation of ECGs is discussed on the basis of 12-lead ECG recording and the morphology of the characteristic waves. The different ECG leads provide different electrical activity of the heart recorded from different orientations. The output of the ECG signal provides information regarding the activity of the heart between positive and negative poles and the way that the heart pumps the blood through the different parts of the body. Different algorithms need to be used in order to analyse the output of ECGs and extract the ECG beat events. In the next chapter, different techniques that have been used for analysis and extraction of ECG beats will be discussed. In addition, some of the feature extraction and classifier techniques will be reviewed as well.

# Chapter 3 Review of ECG analysis and classification algorithms

## 3.1 Introduction

Analysis of ECGs is possible using several different algorithms. Any automated ECG analysis and classification system comprises three main components: pre-processing, feature extraction, and selection and classification of ECG beats. Previous studies have reported the use of different methods for accomplishing each of these tasks. There are different types of classification approaches that can be adopted, such as a LDA classifier, an artificial neural network (ANN) or an SVM classifier. In this chapter, the various techniques that are used for extracting discernible features from an ECG signal will be reviewed, and different pre-processing and classification techniques will be considered.

## 3.2 Pre-processing

The first step, which needs to be considered in ECG analysis, is pre-processing. The pre-processing module usually contains three components, as shown in Fig. 3.1. These are denoising, detection of the QRS complex and extraction of a single beat on the basis of segmentation. Each of these components is described briefly in the following sections.

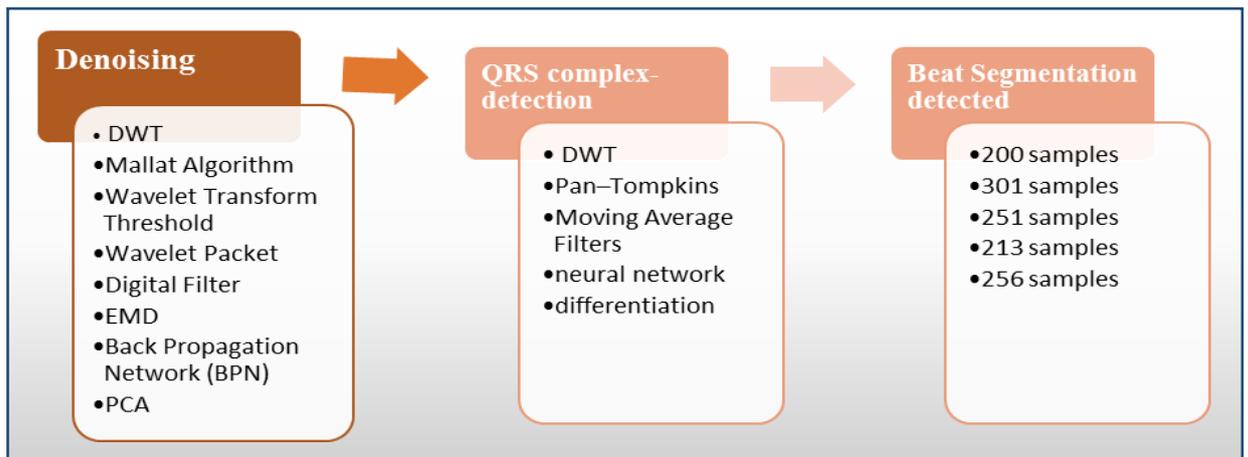


Figure 3.1: The pre-processing module with methods used in previous studies

### 3.2.1 ECG signal denoising

Denoising is the primary processing to remove all the high-frequency noise as well as the power supply interference superimposed on the ECG signal. ECG signals usually contain different types of noise. The sources of these noises can be either cardiac or extra-cardiac. The reasons for cardiac noise are the reduction or disappearance of the isoelectric interval and prolonged repolarisation and atrial flutter, while the sources of extra-cardiac noise are respiration, changes in the electrode's position, muscle contractions and power line interference [52-53]. Moreover, the main sources of ECG noise are the electrical activities of other body muscles, a baseline shift because of respiration, the poor contact of the electrodes, and possible interference from other equipment or electronic devices [54-55].

In 2015, Velayudhan and Peter claimed that there are two main types of noise that affect the ECG signal. The first type of noise is due to the natural features found in the typical ECG signal, these occur predominantly at high frequency and are attributed to electromyogram (EMG) noise, additive white Gaussian noise and power line interference. Noise with a low frequency such as baseline wander are of the second type [56]. Frequency interference, baseline drift, electrode contact noise, polarisation noise, muscle noise, internal amplifier noise and motor artefacts are different types of noises that an ECG signal might contain [57].

Baseline wander comprises the noise artefacts that generally affect ECG signals and it usually comes from respiration; such interference lies between 0.15 and 0.3 Hz [58]. Moreover, since baseline wander is one of the common problems in ECG signal processing, its removal is always required in order to minimise alterations in beat morphology. The baseline wander is removed by using a moving average filter. This can be done easily by using the Wavelet Toolbox [59]. Furthermore, some algorithms have been suggested specifically for removing baseline wander, and these are discussed below.

One approach is through wavelet filtering [60]. Various types of wavelet basis functions (mother wavelet) such as Coiflets, Haar and Daubechies have been used to evaluate the performance of wavelets in denoising ECG signals [58]. Other algorithms include custom-made filter banks, PCA, independent component analysis (ICA) [61], neural networks (NNs),

adaptive filtering [62], empirical mode decomposition (EMD) [63] or wavelet transforms (WT) in conjunction with these [64].

A considerable amount of literature has been published on denoising ECG signals using these various methods. The majority of the publications have used wavelets to denoise ECG signals, remove the 50/60 Hz power line interference [65] and perform other types of denoising [66]. Moreover, ECG signal noise cancellation can be done using adaptive filtering, as indicated in [67]. In [68], notch filter [69], recursive least squares (RLS) adaptive filter [70], normalised least-mean squares (NLMS) adaptive filter [71] and wavelet packet transform [72] were utilised to reduce the 50 Hz power line interference from an ECG signal. An excellent simulation result was achieved when using wavelet packet transform. Furthermore, adaptive filters have been considered to reduce the ECG signal noises arising from power line interface (PLI) and baseline interference. The RLS algorithm is suggested for removing the low-frequency components and tiny features of the ECG [70]. In 2012, a new ECG denoising approach based on alternative noise reduction algorithms such as EMD and discrete wavelet transform (DWT) was proposed to denoise ECG signals, and this method was found to be more effective in reducing noise [73]. In 2014, AlMahamdy and Riley [74] suggested that the most common and essential denoising algorithms are DWT, adaptive filters (LMS and RLS) and Savitzky-Golay (SG) filtering. They applied these methods systematically to real ECG signals in their research entitled ‘Performance Study of Different Denoising Methods for ECG Signals’ and concluded that the wavelet algorithm performed very well.

A recent study by Gang *et al.* (2015) indicated that the three main ECG noises were power frequency interference, baseline drift noise and EMG or muscle artefacts. In their paper, they suggested an improved wavelet packet denoising algorithm. Moreover, they compared the performance of certain other algorithms, such as the WT threshold method and the Mallat algorithm. The wavelet packet algorithm was used to remove the baseline drift and part of the EMG interference in the ECG signal, whereas the method of threshold denoising was suggested and used for the removal of noise from the 60 Hz mains interference and from part of the EMG interference. Using their proposed wavelet packet algorithm produced a better result than that achieved by using the Butterworth digital filter, the Mallat algorithm, the WT threshold method and other standard wavelet packet algorithms [75].

### 3.2.2 QRS complex detection

A QRS wave is associated with ventricular depolarisation. This type of wave has three separate deflections, which are called Q, R and S waves. A variety of noise types and different types of abnormal morphologies are the two main problems that need to be addressed in QRS detection [76]. The R peak location from the patient's recorded data needs to be identified first in order to detect the QRS complex event more accurately in time. The recognition of real peaks on ECG signals is important for diagnosing heart diseases. Various R peak detection techniques are mentioned in the literature. Furthermore, as the precise detection of the QRS complex is difficult, some techniques are recommended to detect the QRS complex.

Detection of the QRS complex plays a central role in ECG classification. Therefore, such a procedure is normally performed prior to feature extraction. There are different methods for QRS detection, such as the Pan Tompkins algorithm [77] or WT [60]. The Pan Tompkins algorithm incorporates several pre-processing steps such as derivative-based peak detection, signal squaring, moving average integration and threshold operations. The logarithm also incorporates linear filtering, a non-linear transformation and a decision rule for the detection of the QRS complex after using a differential operation to find the R peak [77]. A further improvement in this algorithm was developed by Hamilton and Tompkins [78]. The development includes the use of a threshold for peaks and an estimated threshold for noise. A recent study by Krishna proposed a new technique based on the Pan Tompkins algorithm and digital differentiators. The author used infinite-impulse-response (IIR) digital differentiators instead of the original digital differentiator [79]. This not only improved the sharpness of the signal, but also amplified the high-frequency noise in the signal. Careful application of this algorithm is needed as noisy datasets will produce a signal with large number of artefacts when the algorithm is applied.

The last decade has seen rapid development in QRS detection methods. In [4], the QRS complex was detected using multi-rate signal processing and filter banks techniques. Xu and Liu (2004) described an algorithm based on slope vector waveform (SVW) for QRS complex detection and RR interval evaluation. In 2008, Yeh and Wang used an alternative difference operation method (DOM) to detect the QRS complex. The first differentiation of the ECG signal and its Hilbert transform were utilised to find the location of the R peak in the

ECG signal [80]. Izzah *et al.* (2013) indicated that NN pattern recognition is able to classify and recognise the real peaks accordingly. All three waveforms of P, QRS and T were recognised by using NNs. Their study showed that 40 datasets were correctly classified as desired, while nine datasets were misclassified, with a total accuracy rate of 81.6% [81].

Recently, the associated scientific literature around the theme of QRS complexes and R peak detection has been expanded significantly. Some other algorithms have gained popularity. Xia *et al.* (2015) proposed a quick method for the detection of QRS complexes and R peak waves in ECGs using WT and K-means clustering. Real ECG signals from the MIT-BIH Arrhythmia Database (MIT-AD) were used for performance evaluation. In that study, data analysis and pre-processing were done first using the WT. Thereafter, the segmented K-means clustering method was applied to detect the QRS location. A sensitivity of 99.72 and a positive predictive value of 99.80% were obtained as the average R peak detection [82]. Moreover, a new technique based on moving average filters and adaptive thresholding was developed for QRS complex detection. This algorithm was divided into an IIR filter for the pre-processing stages and a moving average filter for the extraction of the QRS complex. Several ECG databases published on the PhysioNet website were used for the testing and evaluation of the technique. Sensitivity between 98% and 99% was obtained when the algorithm was applied to the ECG database [83]. Another new algorithm for QRS detection with minimum pre-processing requirements was presented recently in [84]. To remove the baseline drift, a two-step median filter was applied, while SG smoothing filters (digital smoothing polynomial filters or least squares smoothing filters) were used to smooth the ECG signal. Several standard ECG databases, such as the Fantasia Database (FTD), the MIT-AD, the MIT-BIH Normal Sinus Rhythm Database (MIT-NSD), the BIDMC Congestive Heart Failure Database (CHFD), which is a two-channel database, and a self-recorded dataset were used for the performance evaluation of this algorithm. This algorithm revealed a high accuracy rate of 99.81% for records from the FTD, an accuracy rate of 99.46% for records from the CHFD and an accuracy rate of 99.96% for the self-recorded ECGs. A QRS complex can be detected with high sensitivity on some ECG databases using this new technique. This study also claims that some of the previous studies were successful only on specific databases, whereas the technique proposed in this study has universal applicability. It seems that this study also offers probably the most comprehensive empirical analysis of QRS detection. Furthermore, several other differentiation techniques were considered to create alternative

QRS detection algorithms, such as the differential threshold method and digital differentiators. In [85], an R peak detection algorithm for ECG signals based on the second derivative was proposed, whereas QRS detection using the first derivative was presented by Benitez *et al.* [80]. Lai and Wang (2015) suggested using the differential threshold method for a real-time, complex QRS detection algorithm. The design of this algorithm included three computation steps, an improved differential processor, a detector of the R wave, and a detector of the Q and S waves. In that study, the assessment of the QRS detection using this algorithm was performed using the MIT-BIH databases. Excellent results were obtained for sensitivity and positive prediction of 99.69% and 99.63% respectively [86].

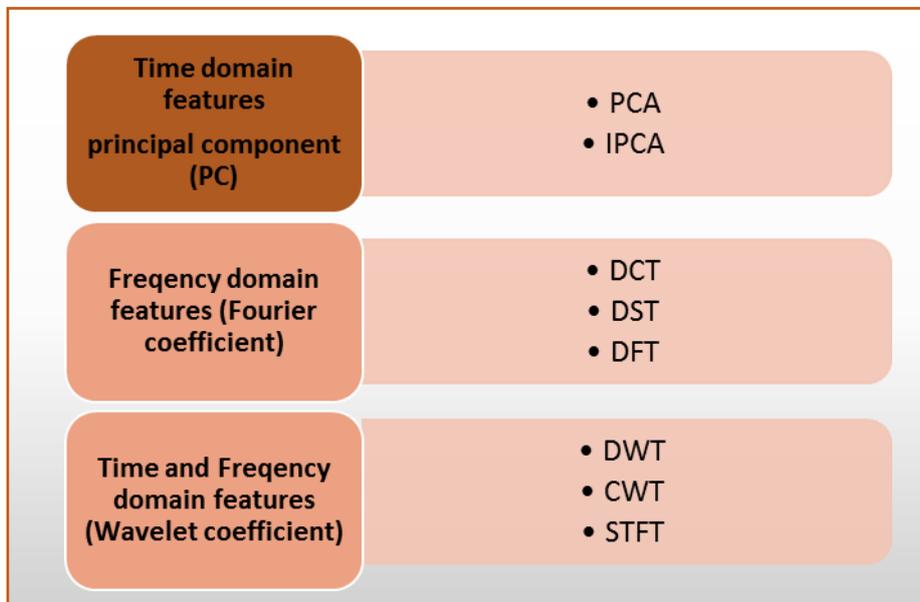
### **3.2.3 Recent advances on beat segmentation**

As stated earlier, each cardiac cycle in an ECG signal consists of the P-QRS-T waves. Segmentation of beats is needed for each cardiac cycle. Segmentation is done after detection of the QRS complex wave. In previous works, different numbers of samples around the R peak were chosen in order to extract the ECG beat. In [87] and [88], a segment or ECG beat of 200 samples was selected. In these works, 99 samples were chosen from the left side of the R peak and 100 samples after the R peak point. In [89], the R peak was taken as the centre of the window used for beat analysis, resulting in 150 samples before the R peak and 150 samples after the R peak. This windowing process sets an ECG beat segment length to be composed of 301 samples, with the ventricular beat placed in the centre of the window. In 2012, Rabee and Barhumi extracted 14 different beat types from the MIT-BIH database using a windowing function with 251 data points. Annotation files from the MIT-BIH were used to allocate the position of the R waves in an asymmetric manner, with 90 points preceding the R wave and another 160 data points after the R wave [90]. Currently, there is no consensus on whether a symmetric window should be used. Since Fourier transforms and WTs are normally used with a unitary apodisation function, such differences in the segmentation procedure are not of consequence to the value of the feature vectors derived through pre-processing.

## **3.3 Feature extraction and selection techniques**

ECG feature extraction plays a significant role in diagnosing most cardiac diseases through ECG classification. Feature extraction and reduction are important procedures that usually influence the performance of any arrhythmia classification system. Therefore, in order to

extract sufficient features and reduce their dimensions to enable optimal classification, a balance between parsimony and feature retention needs to be established. Since each cardiac cycle in an ECG signal consists of the P-QRS-T waves, it is appropriate to consider morphological or timing features such as the widths of the P and the QRS wavelet, the widths of the PQ, PR and QT intervals, the amplitudes of the P and T wavelet and QRS height, including R location as useful features. ECG analysis and feature extraction techniques may thus be classified into three main categories: time domain, frequency domain and time-frequency techniques, as illustrated in Fig. 3.2. In the time domain, the extracted features are the heartbeat interval, the duration parameters of QRS, QT and PR features and the amplitude parameters of the QRS and ST features [91].



**Figure 3.2: ECG feature extraction and dimension reduction techniques**

There are several ECG feature extraction algorithms that have been successfully applied in arrhythmia classification, such as time-domain methods [92], frequency-domain methods [93], time-frequency domain analysis [94], statistical feature signal-based methods [95], fast Fourier transform (FFT), for example, [96] and [97] and Hermite coefficients [98], ICA [99] and PCA [10], Lyapunov exponents [100] and WT, for example, [101] and [102]. Moreover, feature reduction methods often need to be adopted to reduce the dimensions of input feature vectors to the classifier to improve classification performances by improving the generalisation ability of the classifiers. Some of the feature reduction methods, such as PCA and LDA, have also been applied successfully for arrhythmia classification. The PCA method

has been widely used in statistical data analysis, feature extraction, feature reduction and data compression, while the LDA method is an effective supervised dimension-reduction method for pattern recognition problems.

Recently, there have been several studies on extracting different features in ECG signals. Vaneghi *et al.* [103] indicated that the six most commonly used methods are based on autoregressive (AR) approaches, WT, Eigenvector calculations, FFT, linear prediction (LP) algorithms and ICA. Using the Eigenvector method in the frequency domain showed better performance for ECG feature extraction in their study.

WT is one of the preferred methods used to extract features (the wavelet coefficients). These features can faithfully describe each ECG beat segment [30]. According to Asl *et al.* (2008), it is possible to select five features of a total of 15 using a generalised discriminant analysis (GDA) feature reduction method. On the basis of their proposed feature reduction method, they were subsequently able to discriminate between six different types of arrhythmias [104]. In 2007, Übeyli used discrete WTs for feature extraction in conjunction with multiclass SVMs and error-correcting output codes; in Übeyli's work, 265 wavelet coefficients were originally obtained and extracted for each ECG beat. Subsequently, the author used statistical techniques to reduce the dimensionality of the extracted feature vectors. Statistical features (maximum, mean, minimum and standard deviation) of the wavelet coefficients were subsequently used as feature vectors at the input of the SVM classifier [101]. Alternatively, ICA may also be used to reduce the input feature space dimensionally from 200 to 17 features [13]. According to Song *et al.* (2005), LDA and PCA can be used to select a sub-group of only four features from an original number of 17, (which were originally associated with two features correlated to rhythm and 15 features correlated to morphology of the ECG signals) without overwhelming the classifier. Discrete Fourier transform (DFT) is another well-established method for feature extraction. Selected Fourier coefficients can be subsequently used as input to a classifier; this approach has been used extensively in a large number of biomedical applications [17]. In 2011, Gothwal *et al.* utilised an FFT to identify the R peaks in the ECG signal, and NNs were applied to identify different types of heart disease [97]. EMD combined with the Hilbert transform as well as EMD combined with the DFT have also been used to extract the spectral features from ECG raw signals [105]. Moreover, Haque *et al.* (2009) found that using FFT as a feature extraction tool was useful in finding abnormalities in the ECG signal [96].

In recent years, linear and non-linear transformation methods have been applied successfully to ECG signal analysis as well as to the extracted time-frequency-based feature sets. Many studies in the literature also apply two or more techniques in tandem in order to extract and select features of the ECG beats. As a result of using diverse techniques, a combination of various features is used to create the input vector. Some of these studies will be discussed below, as the majority showed good classification accuracy. In 2014, Martis *et al.* proposed a linear method using DWT coefficients featuring a further reduction technique using PCA to extract features from normal and arrhythmia classes. These features were applied to discriminate between normal classes and four arrhythmia classes. Remarkable classification accuracy of 98.78% was shown using an NN for the classification of the five main beat classes [19]. In another study [20], five types of beats were classified with a lower accuracy of 93.48%. The HOS bispectrum in conjunction with PCA was used to capture the features and reduce dimensionality. Furthermore, the feed-forward NN and the LSVM were selected to classify these five types of beats according to these features. Moreover, in the same study, alternative features, such as higher order statistics, morphological features, Fourier transform coefficients and higher order statistics of the wavelet package coefficients, were used in an arrhythmia recognition system. These features were used to classify normal beat (NORMAL), VEB, S, fusion beat (F) and unknown beat (Q) on the basis of a k-nearest neighbour classifier. The overall classification accuracy values were 85.59%, 95.46% and 99.56% for average sensitivity, selectivity and specificity respectively [18]. In 2014, Das and Ari used the S-transform (ST) and WT to select the features of the ECG beat more effectively. Subsequently, a combination of two features was used to create an input vector. Five classes of ECG beats were identified by using these features as input vectors to a multi-layer perceptron neural network (MLPNN) classifier. This method achieved average sensitivity performances of 95.70%, 78.05%, 49.60%, 89.68% and 33.89% for NORMAL, S, F, VEB and Q beats respectively [21]. In another study, Lyapunov exponents, wavelet coefficients and power spectral density (PSD) methods were also used as a set of alternative features and the classifier achieved average accuracy of 93.89% [22]. Morphological and temporal features were combined to classify the five classes of ECG signal, and overall accuracy of 85.9% was obtained using a linear discriminant classifier [23].

Recently, a combination of time-domain and frequency-domain features was used in tandem as input vectors for classification. In 2016, Elhaj *et al.* found that there were some

disadvantages of using a combination of time-domain and frequency-domain features techniques. They concluded that using a combination of linear and non-linear features in the input vector could be an excellent solution. Five types of arrhythmia can be categorised according to five classes with high accuracy (98.91%) when using this technique with an SVM and a radial basis function (RBF) method [88].

In the research in this thesis, some of the methods mentioned above will be used to extract features from the ECG beat on the basis of the Fourier transform, WT or PCA. These methods will be discussed in more detail in Chapter 4.

### **3.4 Literature survey of classification algorithms**

Classification of arrhythmias is a complex problem because of the strict requirement for avoiding false-positive or false-negative results. There are many different approaches that can be used to analyse and classify ECG signals. It has been suggested that LDA [106], back propagation NNs [107], self-organising maps (SOMs) [108], learning vector quantisation (LVQ) schemes [108], SVMs [101] and fuzzy or neuro-fuzzy algorithms should be used [109]. Furthermore, many methods for ECG analysis and arrhythmia detection have been developed in order to increase accuracy and sensitivity; such methods include the use of wavelet coefficients, AR modelling, radial basis function neural networks (RBFNNs), SOMs and fuzzy c-means clustering techniques.

Classification methods can be defined as one of two main types, namely, supervised and unsupervised. Supervised classifiers define a target (class membership) that is used for training; however, unsupervised classifiers have no need for a predefined target or knowledge of the membership. NNs, nearest-neighbour classifiers and SVMs are examples of the supervised classification method, while SOMs and hierarchical clustering are examples of unsupervised classification.

In the following sections, some of the different approaches that have been used in previous studies will be discussed and a comparison of results found in the literature will be provided. Chapter 5 is dedicated to the discussion of other results found in the literature as well as those generated through this study. Results will be placed in context to other signal processing and classification methodologies. Some pattern recognition techniques are also

discussed. Furthermore, an introduction to unsupervised and supervised learning-based approaches within the context of ECG classification will be provided.

### 3.4.1 Supervised classification of cardiac arrhythmias

Supervised learning is based on training a sample from a data source with a target that assigned the correction of each class. Each element in the classifier input space is related to a *pair* value in the output space defined for each class. In other words, the training dataset includes input data and response values. A model can then make predictions of the response values for a new dataset that is built during the training process. Moreover, the supervised learning model is adopted, assuming the presence of the supervisor, which classifies the training examples into classes using the information on the class membership on the basis of each training example. NNs such as feedforward or multi-layer perceptron (MLP) models are typical examples of fuzzy NNs, probabilistic NNs (PNNs) and SVMs [110]. An example of a supervised ECG classification approach is shown in Fig. 3.3 [111].

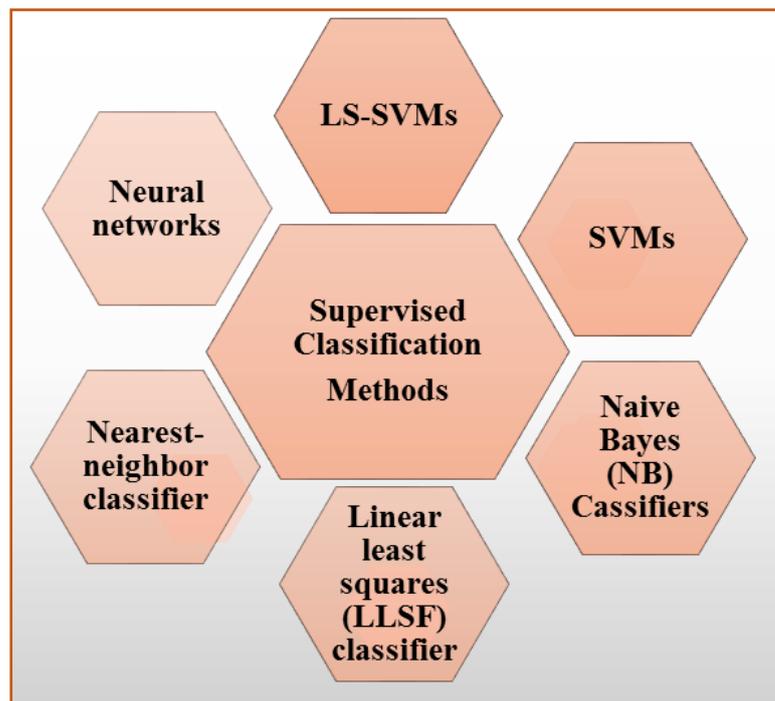


Figure 3.3: Supervised classification algorithm

### 3.4.1.1 Artificial Neural Networks (ANNs)

#### 3.4.1.1.a General structure of an neural network

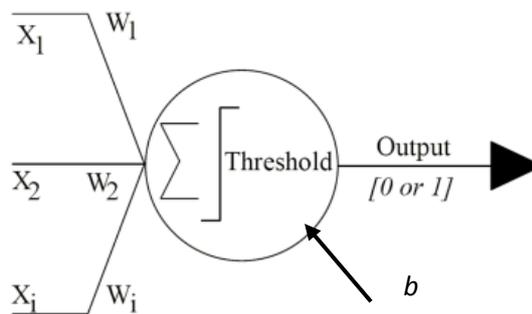
In general, Neural Networks can be divided into three fundamentally different classes of network architectures: Single-Layer Feedforward Networks, Multilayer Feedforward Networks and Recurrent Networks.

#### Neuron Model

The simplest model of neural it contains one input element, one weight matrix, one bias, input to the transfer function  $f$  and an output element. Fig 3.4 shows the structure of the McCulloch and Pitts neurons depicting weights matrix  $w$  where the output of the neuron following thresholding is a zero or one. When a network is represented with an input element, the input matrix and the weight matrix are multiplied and added to the bias. Finally, the output is the function of the total value of this quantity. The output can be written as: [112].

$$O = f(w \cdot x + b) \quad (3.1)$$

where  $x$  input element,  $w$  weight matrix and  $b$  bias.



**Figure 3.4: General description of the McCulloch and Pitts neurons depicting weights and thresholding function**

Example of supervised ANNs include feed-forward and recurrent networks, whereas unsupervised classifiers are often based on self-organising maps.

#### 3.4.1.1.b RBF classifier

The main concept of the RBF classifier is inspired by the theory of function approximation. The classifier is composed of two layer feedforward networks with hidden nodes that implement a set of RBFs. The RBFNNs are non-linear hybrid networks that usually contain a single hidden layer of neurons. Firstly, the input layer broadcasts the coordinates of the input

vector into each of the nodes in the hidden layer. Each node in the hidden layer then produces activation functions based on the associated RBF. Finally, each node in the output layer will compute a linear combination of the activations of hidden nodules [113]. This approach was adopted in a relevant study using non-linear PCA (NLPCA) techniques to classify each ECG segment into normal and abnormal classes [114]. During the algorithm training stages, only normal patterns were used. Only two non-linear features for each ST segment were used. These features were modelled using an RBF network (RBFN) [114].

#### **3.4.1.1.c Bayesian ANN classifier**

The Bayesian ANN classifier is based on a logistic regression model and a back propagation algorithm. A dual threshold method is applied in order to suppress false alarm signals. This classifier acquires arrhythmia properties from the underlying dynamics of the system. It also works even when the dataset includes incomplete information such as missing feature values and unclassified classes. A Bayesian ANN classifier is useful for generating a pattern recognition model based on a given set of inputs and output [115]. The ECG beats are decomposed into finite characteristic waveforms using a sum of Gaussian kernels. A 99.1% accuracy rate for three types of ECG beats (normal, ventricular premature cycle [VPC] and other possible beats) using Bayesian filtering has been reported in [116].

#### **3.4.1.1.d Extended Kalman Filter (EKF)**

The Kalman filter is a state-estimating technique that is implemented to operate on a state-space representation of signals [117]. Meau *et al.* (2006) developed a hybrid system that consisted of the combination of an EKF-based multi-layer perceptron network (MLPN) and a learning fuzzy inference system through the use of a look-up table scheme for the recognition of ECG signals [118]. The Kalman filter method includes a process noise component with different variances that can be estimated on the basis of measurement noise. Kalman filters are based on two approaches: residual-based algorithms or expectation maximisation algorithms [119]. An EKF-based MLPN and a one-pass learning fuzzy inference system can be used to discriminate between various types of abnormal ECG signals. The study by Meau *et al.* (2006) distinguished between four types of ECG signals, namely, VPC, T wave inversion (TINV), ST segment depression (STDP) and supraventricular tachycardia (SVT) from normal sinus rhythm (NSR) ECG signals by using an EKF neuro fuzzy classifier. The ECG waveforms were taken from the MIT-BIH database. Three ECG signals from each

category were taken and used for the training; while a further three different, ECG signals from each category were used for testing after the training was completed. The overall success rate for the detection of STDP was 89.0%, while for the recognition of TINV the success rate was 96.6%. Moreover, for NSR and SVT the classification rates were 88.2% and 99.3% respectively [118].

#### **3.4.1.1.e Fuzzy NN**

Fuzzy networks attempt to model the human reasoning and thinking process. They are composed of generalised crisp sets that have greater flexibility to capture various aspects of the imperfection or completeness of information. Such algorithms operate using fuzzy logic, which is based on fuzzy sets and approximate reasoning. Fuzzy networks can faithfully approximate the inexact nature of the real world found in biological experimental datasets [120]. Moreover, in 2009, Ceylan *et al.* built a module that contained a combination of a type-2 fuzzy c-means (T2FCM) clustering algorithm along with an NN for the classification of ten types of ECG arrhythmias obtained from the MIT-BIH database (normal beat, sinus bradycardia, ventricular tachycardia, sinus arrhythmia, APC, PB, RBBB, LBBB, atrial fibrillation (AF) and atrial flutter) [121]. Furthermore, fuzzy hybrid NN with HOS has been proposed for ECG beat recognition and classification of seven different types of ECG beats. The work indicated that through the recognition of normal beats and the identification of abnormal arrhythmias, different types of beats could be classified to an accuracy of 96% [95]. Moreover, fuzzy adaptive resonance theory mapping (ARTMAP) was selected and used to classify two different conditions from cardiac arrhythmias: normal and abnormal PVC. The results from these tests show that using the fuzzy ARTMAP NN can help classify cardiac arrhythmias with more than 99% specificity and 97% sensitivity [87]. In 2008, Sengur and Turkoglu proposed the use of an artificial immune system (AIS)-based fuzzy K-nearest neighbours (K-NN) algorithm for diagnosing aortic and mitral valve disorders of the heart. The WT and the short time Fourier transform (STFT) were used as feature selection methods. The result showed 95.9% sensitivity and a 96% specificity rate for 215 sample beats taken from the MIT-BIH database [122]. In another study, fuzzy K-NN and NNs in combination with a fuzzy system were used to classify LBBB, RBBB, PVC, and fusion paced and normal heartbeats with classification accuracy of 98% [123].

**3.4.1.1.f Hopfield NN (HNN) classifier**

In order to reduce interference noise, recurrent HNN can be used effectively, as these networks have the additional property that they store information in a dynamic stable pattern. The algorithm from HNN retrieves a pattern that is stored in the memory in order to respond to the presentation of the incomplete or even noisy version of that pattern (equivalent to filtering with *a priori* information). In a study by Bagheri *et al.* (2013), computer simulation results showed that HNNs can effectively model the ECG signal and remove high-frequency noise [124]. Furthermore, Dokur and Olmez (2001) presented a novel hybrid NN structure that can be used for the classification of ECG beats. The method uses feature extractions methods (Fourier and wavelet analysis) for ECG beat classification. Its features are determined by the dynamic programming according to the divergence value of the algorithm. The classification performance, training time and number of hidden nodes for the MLP and the Coulomb energy were also observed in that study. Ten types of ECG beats obtained from the MIT-BIH database were classified using the proposed hybrid NN with a success rate of 96%. The hybrid structure was trained by using genetic algorithms (GAs) in order to increase the classification performance and to reduce the number of nodes [15].

**3.4.1.1.g Recurrent Neural Network (RNN) and Continuous Time Recurrent Neural Networks (CTRNN) classifiers.**

A Recurrent Neural Network (RNN) can be implemented and used as a basis for detection of variability in ECG signals. RNN can be used to classify different types of ECG beats, such as normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat and atrial fibrillation beat that are obtained from different ECG databases. A particular feature of all neural networks is their multi-layered architecture. Multi-layered networks can be classified as feed-forward or feedback networks, according to their connectivity and the direction of information flow. The recurrence allows the network to remember cues from the past without complicating the learning excessively.

An Elman RNN is a network, which in principle is set up as a regular feed-forward network. This means that in this type of network, all neurons in one layer are connected with all neurons in the next layer. Fig 3.5 depicts the architecture of this type of network, it can be seen that the neurons in the context layer (context neurons) hold a copy of the output of the

hidden neurons. Moreover, the output of each hidden neuron is copied into a specific neuron in the context layer.

RNNs can perform highly nonlinear dynamic mappings and thus, have temporally extended applications; whereas multi-layer feed-forward networks are confined to performing static mappings. This is an important advantage over other topologies for the current application, especially when aiming to detect the onset of disease or the impact of lifetime changes to the ECG signal. The strength of all connections between neurons are indicated with a weight, which is similar to a regular feed-forward neural network. At the start point, all weight values are chosen randomly and are then optimized during the training stages. The weights from the hidden layer to the context layer are set to one and are fixed because the values of the context neurons have to be copied exactly in an Elman network. Furthermore, the first output weights of the context neurons are set to be equal to half the output range of the other neurons in the network (for amplitude normalization purposes which leads to better collinearity of the calculated weights). Similar to regular feed-forward neural networks, the Elman network can be trained with gradient descent back propagation and optimization methods [6-7]. Elman networks are, therefore, good candidates for ECG signal classification.

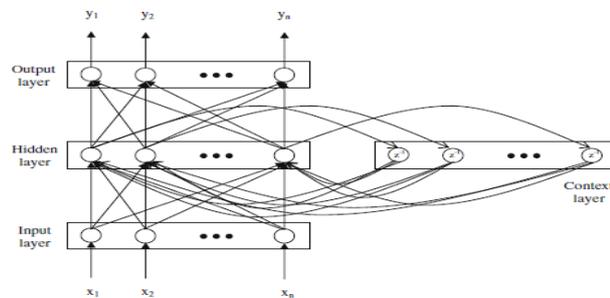


Figure 3.5: Elman recurrent neural network structure adopted from [125].

In recent years, there has also been a growing interest in Continuous Time Recurrent Neural Networks (CTRNN). The popularity of these networks has been increasing because of their simplicity in simulating non-linear dynamical processes. Gallagher *et al.*, (2005) suggested that CTRNNs should be seen as Hopfield type networks with unconstrained connection weight matrices. The unconstrained connectivity provides an improvement in the generalization ability of the networks in the learning process, thus improving its classification ability. Furthermore, CTRNNs are capable of faithfully emulating neuronal activity, and as

such they are a natural platform upon which a classifier can be built for ECG diagnostics. CTRNNs are made up of neurons and each neuron's activity can be described by the following expression:

$$\tau_i \frac{dy_i}{dt_i} = -y_i + \sum_{j=1}^N w_{ji} \sigma(y_j + \theta_j) + s_i I_i(t), \quad i = 1 \dots N \quad (3.2)$$

In the above expression,  $y_i$  is the internal state of neuron  $i$ ,  $\tau_i$  is a time constant of neuron  $i$ ,  $N$  is the total number of neurons,  $w_{ji}$  is the strength of the connection from neuron  $j^{th}$  to neuron  $i^{th}$ ,  $\theta_j$  is a threshold/bias term,  $\sigma(x) = 1 / (1 + e^{-x})$  is the standard non-linear (sigmoid) logistic activation function and  $I_i(t)$  represents a weighted sensory input with strength  $s_i$  [126].

The main difference of a CTRNN over other kinds of neural networks is that the neurons could propagate a signal back through the network. Other neural networks considered in ECG are feed forward, which means that neuron signals could only be unidirectional. In addition, CTRNNs are more dynamic in terms of mimicking biological neuronal signal discharge processes. Moreover, CTRNNs are also deemed to be more efficient than other neural networks in terms of computations since they could be used to directly simulate each spike [127]. The CTRNN has additional advantages and computational efficiency over other discrete formulations. For example, using a discrete-time RNNs there is a considerable dependence of the resulting models on the sampling period used in the process, whereas for CTRNNs this can be varied without the need for re-training. Even in the presence of measurements noise the RNNs are capable of providing long-range predictions. Another advantage of CTRNNs is that if they are compared with other NN types such as feed forward neural networks (FFNNs), they have been shown to be more efficient in terms of the number of neurons required to model a dynamic system [128].

An RNN was implemented and used as a basis for the detection of variability in the ECG signals in [6]. Four types of ECG beats were considered: normal, congestive heart failure, ventricular tachyarrhythmia and atrial fibrillation beats, these were successfully classified using an RNN trained with the Levenberg–Marquardt algorithm [6].

#### **3.4.1.1.h ECG Classification using other neural network types**

Added to the above there is also a plethora of other alternative algorithms that have also reported very high classification accuracy. Guler and Übeyli (2005) showed that four types of ECG beats (normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat, and atrial fibrillation beat) were successfully classified with a 96.94% rate of accuracy by using Combined Neural Networks (CNNs) trained with the extracted features from discrete wavelet transforms of normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat, and atrial fibrillation beat signals [129].

PNN are another alternative modality for ECG classification. A number of types of arrhythmias have been classified by using these neural networks. By combining PCA with LDA, and a PNN classifier it was possible to categorise eight different types of arrhythmia from ECG beats [10].

Furthermore, BPNN and feed-forward neural networks are among the most widely used groups of arrhythmia condition classifiers. A discrete wavelet transform (DWT) approach and a multilayer feed-forward neural network with one layer of hidden units were used to classify 13 types of ECG beats, with a 96.77% overall accuracy of classification. A set of DWT coefficients, which contain the maximum information about the rhythm, was selected from 4-level wavelet decomposition. For training a feed-forward neural network back-propagation with momentum was used [130]. Discrete wavelet transform and Back Propagation Neural Network (BPNN) were used for binary (normal versus abnormal) ECG detection. A 96% rate of accuracy is reported for the normal class sample, whereas a 100% accuracy rate is achieved for the abnormal class sample. Moreover, 97.8% of accuracy was obtained for the overall system using a BPNN classifier [58]. In another study, a multilayer feed forward neural network with a back propagation algorithm is employed to classify both normal and arrhythmia spectra. The three layer networks were shown to be sufficient to approximate any nonlinear function that can be encountered in ECG classification studies [4].

A Multilayer perceptron neural network (MLPNN) was also adopted in a study by Übeyli and Güler (2004) to classify the normal beats and partial epilepsy beats with a 97.50% rate of accuracy; this was also a study that used the Levenberg–Marquardt algorithm for training the MLPNN [131]. In another study, Güler and Übeyli (2005) demonstrated the usefulness of MLPNN algorithms, when trained by various means e.g. backpropagation (BP),

delta-bar-delta (DBD), extended delta-bar-delta (EDBD) and quick propagation (QP) approaches. They showed that using MLPNN trained with these algorithm can help classify normal beats and partial epilepsy beats with an accuracy of 90.63%, 92.50%, 95.00% and 96.88% respectively [132]. By comparing the reported results in these studies, it seems that the use of the Levenberg–Marquardt algorithm for training can lead to slightly higher percentage accuracy than using a simple back propagation algorithm.

The evidence presented in this discussion suggests that different types of NNs can be used for creating a successful classification module. A SVM, or support vector network, is another supervised machine learning technique, which has been used for ECG classification. Over the past decade, a significant body of literature in ECG analysis has focused on binary and multi-class SVM classification. The following section will discuss some of these studies, as this was one of the preferred directions to pursue further research in the current project.

#### **3.4.1.2 Support Vector Machine (SVM) and kernel functions**

Application of SVM based classification represents the cutting edge of current research in pattern recognition. These innovations include finding a hyperplane with the widest margin between the two classes and using a higher dimensional setting through a kernel function. These two objectives can be formulated using a Quadratic Programming (QP) framework. This makes SVMs an effective pattern recognition and classification solution platform in various fields such as bioinformatics (e.g for gene expression) and physiology (e.g lung data classification). The process of analyzing ECG signals and classifying arrhythmia problems through SVM, especially in detecting QRS complexes for ECG beats classification is well established [133]. In that work, a 12-lead ECG recording were studied and T, P, and QRS complexes were detected using an SVM classifier. After filtering of disturbances in the ECG signal, the training data is formatted in such a way that the LIBSVM software can be used. The training instance matrix is a  $m \times n$  matrix, where  $n = 12$  represent the number of ECG leads used. After the SVM training, each record is tested for the recognition of QRS complexes. The detected QRS complexes are removed from the signal of the ECG and replaced by a baseline. The ECG signal is then processed for the recognition of T-waves. Calculation and normalization of the ECG signal gradient, SVM training, and testing are three important preliminary steps before detecting the T waves. The T waves are firstly removed in the ECG signal without the QRS complexes and replaced by a baseline. The ECG signal is

then processed for the detection of P waves. LIBSVM software is used for the detection of these waves. This is a software package that is widely used for software vector classification and regression, as well as parameter estimation [134].

In 2013, Sambhu and Umesh used Wavelet Transform and Support Vector Machines to classify seven types of heartbeats. The classification was implemented by using OAO (One-Against-One) SVMs and the result indicated an accuracy of more than 97 % for all classes. The features that were selected were also related to a particular type of disease; because the single feature vectors were mixed in a matrix that included three types of features, namely temporal features, morphological features and statistical features. Temporal features such as ST Intervals, RR Intervals, Heart Rates, TT Intervals, PR Intervals, QT Intervals and PP Intervals were included in the findings, while the statistical features consisted of the mean, variance, skewness, kurtosis, sum, root mean square and mean absolute deviation of the wavelet coefficients. The authors of the work conducted multiclass SVM classifications, using the One-Against-One (OAO) approach, and linear kernel functions [135].

In 2005, Zhao and Zhang published a study [136] on the use of wavelet transforms and SVM for ECG feature extraction and classification. The result of their study attained a 99.68% accuracy of classification for the recognition of six heart rhythm types, which was an improvement of the result indicated in [135]. For detecting the R peak, the Pan and Tompkins algorithm was used, with four decomposition levels. It was indicated that the detail wavelet coefficient  $d_1$  is usually a noise signal that must be removed, while  $d_2$ ,  $d_3$  and  $d_4$  denote higher frequency components of the ECG signal. Furthermore, they showed that the approximation wavelet coefficients at the fourth scale represent the main feature of each heartbeat. In order to obtain optimal SVM classification, two parameters SVM parameters  $C$  and kernel parameter  $\gamma$  should be used. In addition, the best value of these parameters  $C$  and  $\gamma$  could be found by using the cross-validation method. Martis *et al.*, (2012) indicated that through PCA reduction of time domain signals, feed forward NN and LS-SVM can be used to classify five types of ECG beats (NORMAL, RBBB, LBBB, APC and VPC) with an accuracy of 98.11% [137]. In another study, fast LS-SVMs classification and DCT were used to classify six types of ECG beats (normal beat, LBBB, congestive heart failure beat, PVC beat, non-conducted P-wave, VEB) with a 95.2% rate of accuracy [24]. Rabee and Barhumi (2012) used SVM with wavelet multi-resolution (through the discrete wavelet transform) pre-processing for classifying 14 different types of heart beats, the work was based on a comprehensive study

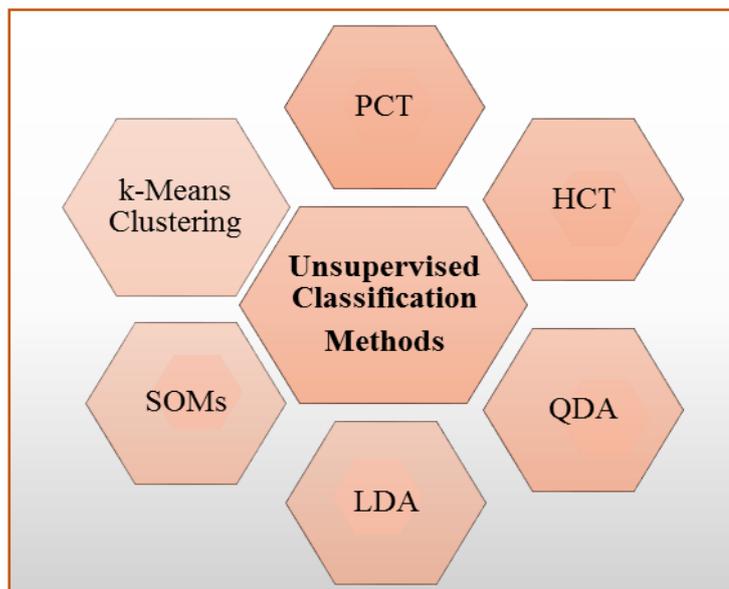
that took into consideration, 17260 ECG beats, selected from the MIT/BIH database. The average classification accuracy achieved was 99.2% [90]. Moreover, extracted principal components of the bi-spectrum using an LS-SVM with an RBF kernel were used to classify the same five types of beats with a 93.48% rate of accuracy [20]. In a further study, Übeyli (2007) classified these signals using multiclass SVM and discrete wavelet transforms (DWT). Two different studies with different feature vectors were conducted. In the first study, 265 wavelet coefficients were used, and the total classification accuracy was 95.56%. In the second study, statistical features such as maximum, mean, minimum and standard deviation of the wavelet coefficients were used to create the input feature vectors. Reducing the dimension of the feature vectors showed a significant increase in classification accuracy to 98.61% [101]. In studies by Martis *et al.*, (2013) five types of beat classes of arrhythmia were analysed, namely non-ectopic beats, supra-ventricular ectopic beats, ventricular ectopic beats, fusion beats and unclassifiable and PB. It was found that the DWT could provide acceptable timing and good frequency resolutions for ECG datasets. In their study, a SVM, neural network (NN) and PNN were used for classifying these five types of ECG beats for automated diagnosis [87]. One should be very critical to that study, however because of issues related to Nyquist sub-sampling.

Generally, as a conclusion to the presented literature review in the previous pages Machine Learning techniques should be strongly recommended for classifying ECG signals because of their superior generalization capability compared to other traditional classification techniques. Machine Learning techniques are also characterized by higher classification accuracies and lower sensitivity to high dimensionality in the dataset. They can therefore be used to also enhance other learning techniques such as support vector machines by significantly boosting the latter's generalization capability. They can also enhance SVMs' robustness against the problem of limited availability of training beats [138].

### **3.4.2 Unsupervised classification of cardiac arrhythmias**

With unsupervised classifiers, there is no need for a target or knowledge of the membership function of a sample. Indeed, when it comes to unsupervised learning models, the pattern class information and reinforcement learning are identified and learnt through trial and error interactions with the environment. This approach implies an ability to learn and organise information without providing target information to evaluate the potential solution [110]. Fig.

3.6 illustrates some widely used models of unsupervised classification methods, including partitioning clustering techniques (PCT), hierarchical clustering techniques (HCT) and quadratic discriminant analysis (QDA) [111].



**Figure 3.6: Unsupervised classification algorithm.**

SOMs [110] are one of the examples of unsupervised classification that have been used to classify cardiac arrhythmia conditions. SOMs are applied for classifying different types of ECG beats such as normal and abnormal beats. As explained in the above study, they have been used to design a customised heartbeat classifier and a global beat classifier. Afterwards, the two classifiers may be combined through the mixture-of-experts principle. In the study by Wen *et al.* (2009), they used a self-organising cerebellar model articulation controller (SOCMAC) network. The SOCMAC algorithm represents the combined structure of the CMAC network into the Kohonen layer of the SOM. This approach can effectively distribute the learning error into the memory contents of hyper-cubes [117]. Talbi and Charef (2008) introduced another approach using SOMs that are combined with information from the QRS complex. The combination of SOMs and QRS complexes is used to discriminate PVC arrhythmia through the use of the power spectrum density of the QRS complexes [139].

In 2004, Gaetz *et al.* proposed using SOMs to analyse the ECG signals of patients suffering from depression. The 84 subjects were divided into groups. Overall, the correct classification was very low, however, between 54% and 70.2% [140]. It is worth noting that there are more than 15 types of ECG beats that can be of interest to physiologists. Therefore, the generalisation ability of the classifier as well as its tuning parameters need to be studied

extensively before they can be used reliably to address specific heart conditions. Such an approach would also provide wider acceptance of other approaches such as that of Lagerholm *et al.* (2000), who suggested that 16 ECG beat types can be classified using a Hermite function decomposition coupled with an SOM to an accuracy rate of 98.5% [141]. Hermite basis functions and SOM have been used successfully for clustering ECG beats. Firstly, each QRS complex taken from the MIT-BIH-AD was reduced into Hermite basis functions in order to find coefficients and width parameters that are used to represent the QRS complex. Self-organising NNs were then applied to cluster the data into 25 groups [139].

Hierarchical clustering is another unsupervised learning method that has been proposed for the analysis of large sets of ECG data classified into groups, as discussed by Nishizawa *et al.* Firstly, this algorithm arranges the data in a multidimensional space without any pre-processing and then classifies it into clinical groups according to the criteria defined by the user. The classification result was represented via a binary tree structure as a hierarchy of clusters, and the simulation results indicated a good performance. Another interesting study was conducted using the unsupervised classification of QRS complexes through the use of a hierarchical cluster analysis and a two-step correlation technique. The ECG signal was first filtered by means of two median filters in order to reduce baseline wander and high-frequency noise. Three different feature sets were used for cluster analysis: four Fourier coefficients, the first seven coefficients of a Hermite series, and the first eight discrete cosine transform (DCT) coefficients [142].

K-means clustering methods based on integrated shape averaging (ISA) were proposed to characterise P wave shapes and to distinguish between patients at risk of AF and non-patients [143]. The results obtained show that an unsupervised learning method was able to differentiate clinically between subgroups in a meaningful way using ECG information. More research needs to be conducted in order to improve the classification accuracy of unsupervised learning algorithms, however. Furthermore, such algorithms are more difficult to validate and as a consequence may not be adopted for routine use in hospitals in the near future.

### **3.5 Summary**

In order to fully interpret the ECG records, pre-processing, segmentation, feature extraction and selection, as well as classification must be performed. This chapter reviewed several techniques for QRS complex detection and heartbeat classification for the diagnosis of common pathologies. Both supervised and unsupervised classifier techniques that have been used for ECG beat classification such as neural network and SVM were reviewed and discussed. In the following chapter, feature dimensions and classifier parameters are altered in a systematic manner to optimize them. The analysis is performed in the time or frequency domains.

# **Chapter 4 Denoising and classification algorithms using CSVM and SIMCA**

## **4.1 Introduction**

As stated in the previous chapter, correct ECG beat classification requires noise removal and QRS complex detection, feature extraction and selection, and signal classification. The classification output can be used for making a report of the patient's heart condition. In the following section, the different algorithms for noise-free parsimonious feature extraction to generate the input vector to the classifier will be presented. For frequency domain feature extraction, discrete cosine transforms (DCTs), discrete sine transforms (DSTs), and the discrete Fourier transform (DFT) are used. This chapter also discusses the algorithms and technical aspects of some important classification algorithms such as CSVM, MSVM, and SIMCA, which will be used to classify ECG datasets, as discussed in a subsequent chapter (chapter 5).

## **4.2 Data collection**

In this study, data are selected from three different databases: the MIT-BIH Arrhythmias database, the European ST-T Database, and the St Petersburg INCART 12-lead Arrhythmia Database (incartdb). Normal beat, and some abnormal beat types, are extracted from these databases for a number of patient records. The durations of beat segments varied between experiments, for example by using windows of either 301 or 256 data points, as indicated in a number of previous studies [24],[89]. Therefore, the duration of waves and beats depends on the type of lead used to record condition settings, as well as the pathology of the arrhythmia conditions. Different conditions are investigated through different length datasets.

## **4.3 Signal processing software tools used**

### **4.3.1 MATLAB 2012a software and associated tools box used in the current study**

MATLAB is a powerful and comprehensive software that can be used for signal processing and medical imaging. It can also be used to perform data classification. It is a high level language for technical computing. As an interactive system, it entails computation,

visualisation, and programming, expressed in mathematical notations. This includes mathematical computations, algorithm development, modelling, simulation, data analysis, exploration and visualisation, scientific and engineering graphics, and application development. MATLAB, which stands for “*matrix laboratory*”, has evolved to become user-friendly for a variety of consumers, including universities and industry.

Furthermore, MATLAB 2012a contains several toolboxes and there are various uses for these that make numerical processing easier. These are generalised to perform the following functions: numerical integration (quad), discrete Fourier transform (fft, ifft), statistics (mean, median, std, var), curve fitting (cftool), signal processing (sptool), and numerical integration of systems of ODEs (ode-45). Assuming an advanced level of understanding by the prospective user, the signal processing toolbox, the wavelet toolbox, the library support vector machines (LIBSVM), and the WFDB software package are all adopted in the current study [144]. These enabled programming and scripting of the required mathematical expressions discussed in this chapter. A brief discussion of these toolboxes follows.

#### **4.3.1.1 Signal processing toolbox**

This toolbox is a collection of automated routines that are built on the MATLAB numeric computing environment. It contains processing operations in the following two categories: signal processing command line functions and a suite of graphical user interfaces for interactive filter design, signal plotting and analysis, spectral analysis, filtering signals, and analysing filter designs. Its main functions are algorithms, which are used for expressing M-files and implementing signal-processing tasks. As a specialised extension for the MATLAB computational and graphical environment, it was created for the benefit of computation and visualisation. Its basic entities are signals, systems, and functions, such as digital or discrete signals and filters [145]. Some functions from this toolbox were used in order to extract features in the frequency domain using Fourier transforms such as DCT and DFT.

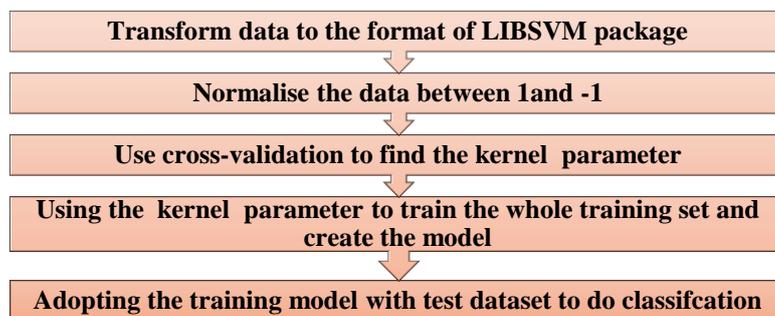
#### **4.3.1.2 The MATLAB 2012a wavelet toolbox**

This toolbox is a collection of functions and applications that analyse and synthesise signals and images as well as data that exhibit regular behaviour punctuated with sudden changes. This toolbox also includes algorithms for continuous wavelet transforms (CWTs), scalograms,

and wavelet coherence. The toolbox enables the user to analyse the frequency of content and signal changes to reveal time-varying patterns found in multiple signals. Because of this, multiresolution analysis from fine-scale to large-scale features can be done as well as identification of discontinuities and detection of change points that are not visible in unprocessed datasets. It also features compression and reconstruction of signals and images to match pursuit algorithms and lifting methods for constructing custom wavelets [59]. Some functions from this toolbox were used in this thesis in order to do the DWTs and find the wavelet coefficients to be used as input to the classifier. In addition, the DWT was used to extract features in the time-frequency domains.

#### **4.3.1.3 A Library for Support Vector Machines (LIBSVM)**

This is an integrated software that supports vector classification, regression, and distribution estimation. The main features of LIBSVM include different SVM formulations, efficient multi-class classification, probability estimates, and various kernels, enabling weighted SVM for unbalanced data. Furthermore, it is compatible with both C++ and Java resources as well as various interfaces, codes, and extensions and provides automatic model selection that generates a contour of cross-validation accuracy [146]. It also features an easy script (easy.py) that automates everything from data scaling to parameter selection. A typical use of LIBSVM entails two stages: training a data set to obtain a model and then using this model to predict information of a testing data set and performing classification. The LIBSVM software was used for classification and implementation of SVM classification. It enabled the simple scaling on the data, the choice of the kernel functions and cross-validation to find the best parameter value, adoption of the best parameter value to train the whole training set and it was finally used to generate the model to be used with the test data [147]. Fig. 4.1 summarises the main procedure that must be followed for performing classification using the LIBSVM package.



**Figure 4.1: Procedure of LIBSVM classifiers.**

#### 4.3.1.4 Wave Form Database (WFDB ) Software Package and toolbox for MATLAB 2012a

As a tool for PhysioNet, the WFDB software package is a compilation of the WFDB library. It is a separate shared library supported by a platform. It uses Java Native Interface (JNI), with the WFDB library being coded in Java. Its major components include the WFDB library, the WFDB applications for signal processing and automated analysis, and the WAVE software for viewing, annotation, and interactive analysis of waveform data. Documentation, featuring tutorials and reference manuals, is also included in the package. The WFDB toolbox for MATLAB 2012a includes over 30 functions and utilities that work with PhysioNet [27]. PhysioNet is an open-source project support software, comprising applications and medical databases [148]. To find the R peak location in the QRS complex, several functions from this toolbox were used including (rdann), used to read annotation files for patients records, and (rdsamp), used to read signal files of records.

#### 4.4 General classification methodology in separate stages

The block diagram of the proposed method for ECG beat classification is described in Fig. 4.2. As can be seen, the method is divided into three steps: ECG sampling and pre-processing, calculation of the feature vector used as the input to the classifier, and classification. These methods are described separately in this section using a variety of algorithms to achieve each of these steps.

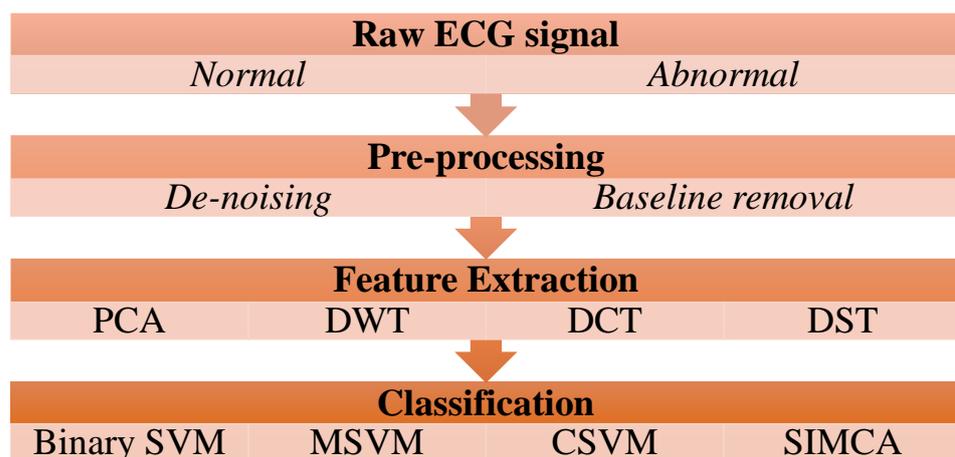


Figure 4.2: Structure and methodology of ECG classification.

### 4.4.1 Signal pre-processing

Pre-processing must be performed further before conducting analysis of the signal. This step usually involves the use of different filters for denoising purposes. Removal of noise from the ECG signal helps to improve peak detection and beat feature extraction. In its simplest form, a digital filter may be implemented, using the following general difference equation to remove noise from the sampled ECG signal:

$$x_d(n) = x(n) - Ax(n-1) \quad (4.1)$$

where  $x(n)$  is the original input ECG signal at time instance  $n$ , which is used as an input, and  $x_d$  is the output signal at time  $n$ .

Different types of noise such as electrode motion and electrical lead electronic noise, power-line interference, baseline wander, and muscular movement can corrupt the ECG record [149].

#### 4.4.1.1 Baseline removal and de-noising

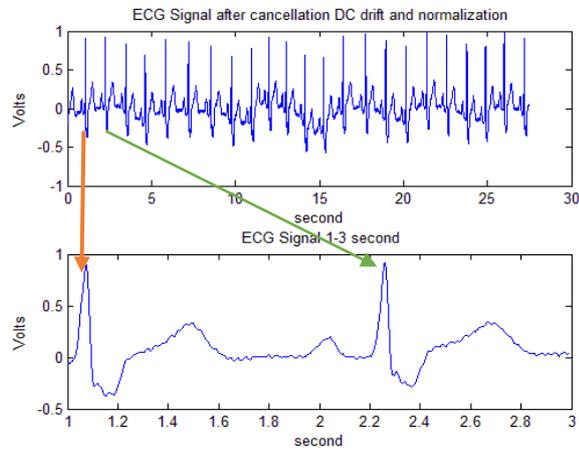
Baseline wandering is attributed to the respiration process and lies between 0.15 and 0.3 Hz [58]. Baseline wander removal is one of the most common problems in ECG signal processing, and its removal is required in order to minimise possible changes in beat morphology. The baseline wander of an ECG waveform is removed by first loading the original signal, and then smoothing it in amplitude  $y$  using a moving average; this can be conveniently accomplished using the Wavelet Toolbox [59].

Denoising is also necessary to remove high frequency, as well as power supply interference from the ECG signal. Several research groups have been consistently using wavelets for denoising of biomedical signals [60]. Different approaches are well established after adopting wavelet basis functions, such as Coiflets, Haar, and Daubechies [58].

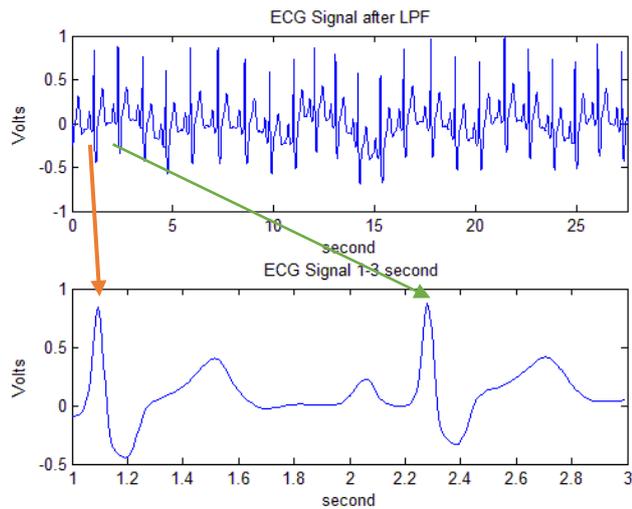
#### 4.4.1.2 R peak detection algorithms in the QRS complex wave and extraction of heartbeats

In the last few decades, various algorithms have been introduced for detecting specific features in the ECG wave. Moreover, several techniques have been developed in order to improve the accuracy of the QRS complex detection. Such approaches use sophisticated tools such as neural network [150] or fuzzy hybrid neural networks [95]. As stated in chapter 3, Pan

and Tompkins [77] established one of the most popular methods for the detection of the QRS complex. This algorithm, based on linear filtering, is followed by a non-linear transformation and the application of a decision rule algorithm. The diagram below illustrates the detection of QRS complex using the Pan-Tompkins algorithm, after cancellation of DC drift and normalisation, low pass filtering, high pass filtering, derivative filtering, squaring, and moving window integration. Fig.4.4 shows the systematic progression through all of the different stages, so that a visual understanding of the algorithm can be developed (the result of low pass filtering). Fig. 4.5 provides the result of high pass filtering. The results of derivative filtering and squaring are illustrated in Fig 4.6 and 4.7, respectively. In all plots, a magnification of the data to include only two beats is provided.



**Figure 4.3: Cancellation DC Drift and normalisation.**



**Figure 4.4: ECG Signal after low pass filter (LPF).**

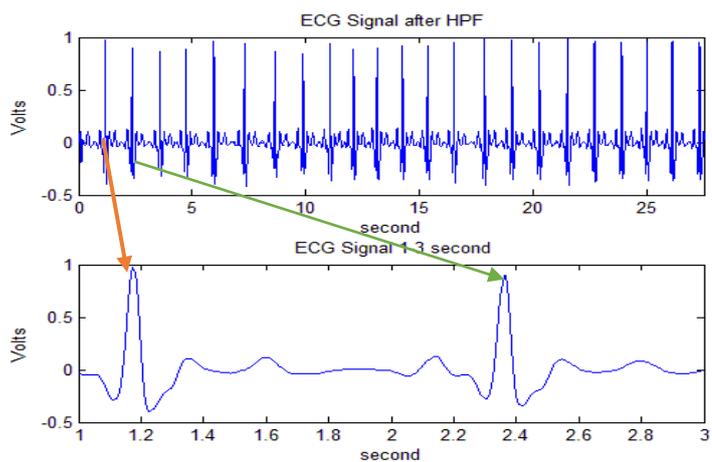


Figure 4.5: ECG Signal after high pass filter (HPF).

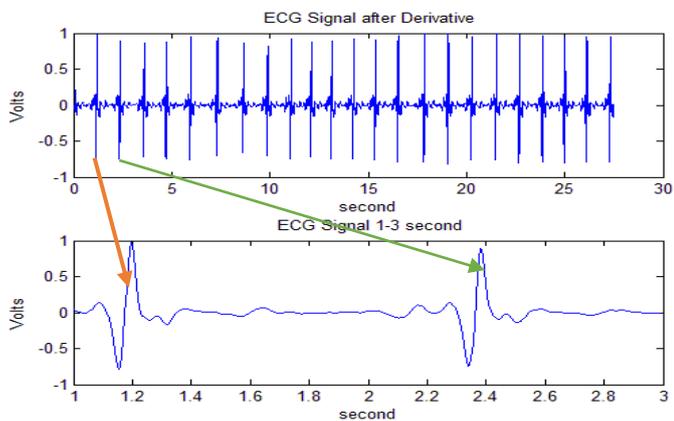


Figure 4.6: ECG signal After Derivative.

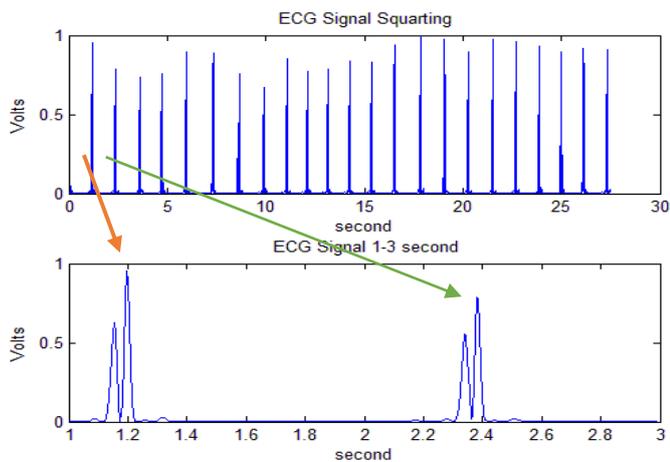
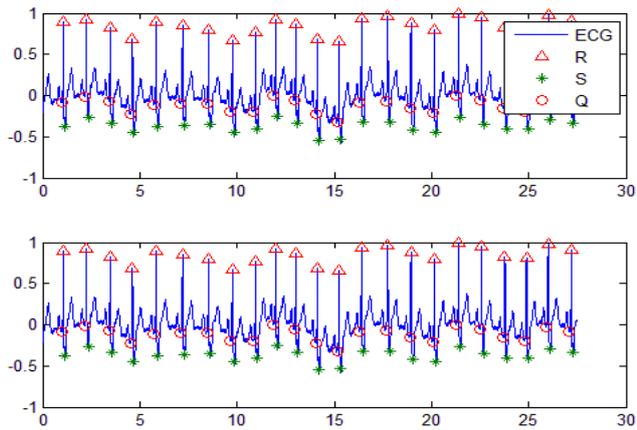


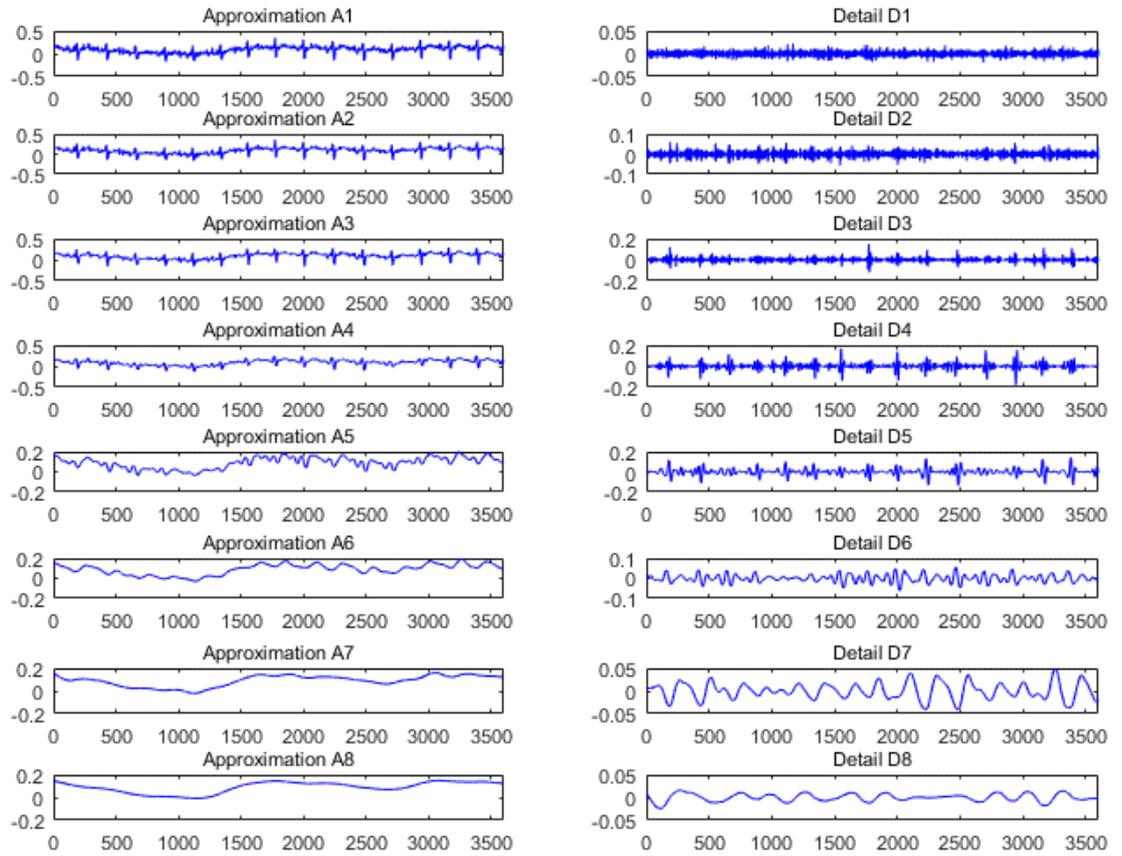
Figure 4.7: ECG signal Squaring.



**Figure 4.8: QRS complex Detection.**

Another method that has been used for extracting an ECG wave involves the use of multiresolution wavelets and filter banks. In 2010, Pal and Mitra proposed using a multiresolution wavelet transform for the detection and evaluation of the QRS complex, as well as P and T waves. The original 12-lead ECG recordings were collected from the Physionet PTB diagnostic database, in order to perform and validate the proposed algorithm. Over 99% detection accuracy for R peak was achieved. Furthermore, detection accuracy of heart rate, P wave, QRS complex, and T wave were over 97%, 96%, 95%, and 98%, respectively [151].

In this study, the multiresolution wavelet decomposition [152] extracted features in both time and frequency domains. The algorithm was applied on selected ECG data records no 209, taken from the MITBIH Arrhythmia Database. Decomposition of the signal was performed up to level eight, as shown in Fig 4.9. Fig 4.9 illustrates reconstruction of approximation and details coefficients at each level.



**Figure 4.9: Reconstruction of detail and approximation coefficients at eight level of decomposition.**

It can be noted that small scales represent the high frequency components, while large scales represents the low frequency components of the signals. Detail coefficients ( $d$ ) at level 3, 4, and 5 are recognised for the detection of  $R$  peak using the following expression.

$$S1 = d_3 + d_4 + d_5 \quad (4.2)$$

$$S2 = \frac{d_4 \times (d_3 + d_5)}{2^n} \quad (4.3)$$

$$S3 = S1 \times S2 \quad (4.4)$$

where  $n$  is the level of decomposition.

Following the  $R$  peak detection, the  $Q$  and  $S$  waves are then identified using the expression below:

$$S4 = d_2 + d_3 + d_4 + d_5 \quad (4.5)$$

The  $Q$  and  $S$  points are identified from the first zero slope points on each side of the  $R$  peak.

The following formula is used to perform five-point differentiation on  $S4$ :

$$f'(x) \approx \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} \quad (4.6)$$

where  $h$  is the time step.

For the detection of the T and P waves, reconstruction detail coefficients at levels 6 and 7 are used according to equation (4.7), as it is believed that the energy of both T and P waves are mostly at scale levels 6 and 7. This is an assumption that requires reassessment on the basis of previous work performed at Reading by Froese *et al* [8]. The peak in the T wave is recognised as the maximum after detection of the S point. In addition, the onset and offset of the T wave are detected as the minimum estimated crossing points on each side of the T peak [151]. The expression for finding T peak is:

$$S5 = d_6 + d_7 \quad (4.7)$$

Alternatively, the QRS complex can be found using multi-rate signal processing and filter banks, as indicated in [153]. Some QRS complex detection algorithms are based on first finding the R peak location, before finding the other waves. In [154], an algebraic approach based on differential algebra and operational calculus was used to find the R peak location in the ECG signal. In addition, the first differentiation of ECG signal with its Hilbert transform is used to find the location of the R peak in the ECG signal [80].

In this study, the R peak location in the QRS complex was identified using the WFDB Software Package. Following identification of the R peak, ECG beats were extracted after inclusion of 150 samples before the R peak with a window spanning 150 data points after the R peak, as shown in Fig. 4.10a. The majority of studies reviewed in this project have used this technique for extracting ECG beats from each database.

#### 4.4.2 Feature extraction and selection

As stated earlier, the ECG features can be extracted in either the time [92] or the frequency domain [155]. The compression of the ECG signal can be performed directly by using time domain techniques. Intelligent sample selection criteria must be used in order to get a high performance time domain compression and to extract the important features in the signal. An inverse process is used to reconstruct the original signal with a few features [156]. Frequency domain techniques, based on dividing the signal into frequency components and allocating

ECG bits in the frequency domain, are also possible. The input ECG signal is distributed into blocks of data and stored as frequency domain features in the form of a vector [157]. Transforms, particularly integral transforms, are usually utilised to reduce the complexity of solving mathematical problems. The differential and integral equations, through judicious application of the opposite transforms, may be changed into algebraic equations where the solutions are more easily acquired using numerical techniques. The Fourier transform helps decompose a signal into its frequency components, while the Karhunen-Loève transform (KLT), decorrelates a signal sequence. These are the most commonly used transforms in signal processing. In exploring the DCT, it is important to consider the continuous versions, such as the Fourier cosine transform and the Fourier sine transform. The properties of continuous transforms are well known and bear great resemblance to those of the DCT and the DST [158]. In the following section, some of the most well established ECG time and frequency domain feature extraction algorithms are discussed, including Direct Time-Domain Techniques and Transformed Frequency Domain Techniques (such as DFT, DST, and DFT). The coefficients of these transforms comprise the input vector to the classifier.

#### 4.4.2.1 Discrete Fourier Transform (DFT)

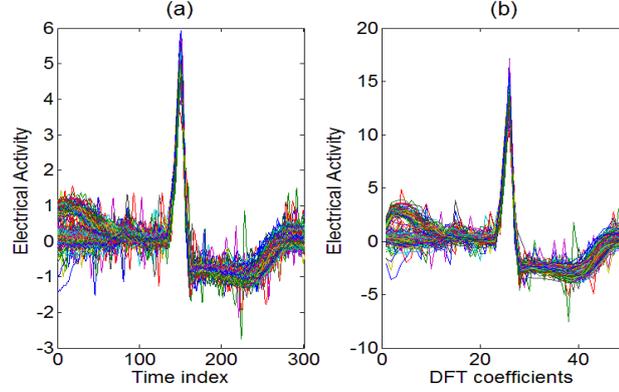
The Fast Fourier Transform (FFT) is an efficient algorithm which has been used to compute the discrete Fourier transform (DFT) and its inverse. The DFT has been used to identify and extract features from ECG beat signals in the frequency domain. The selected coefficients form the elements of a row vector that defines the classifier input feature space. DFT coefficients were used to create the training as well as the testing datasets used for ECG classification. Initially, the DFT is applied to the ECG beat extracted using Eq. (4.8a). In Eq. (4.8a),  $x(n)$  denotes the input signal amplitude in the time domain, while  $X(k)$  the frequency domain output expression. Eq. (4.8a) is used to find a row vector for DFT coefficients for each beat using the DFT matrix, whereas (4.8b) is used to reconstruct the original signal from fewer DFT coefficients. DFT coefficients are transformed back to the time domain using an inverse DFT. From the reconstructed segments, threshold data points are eliminated to generate a signal with only few features, as shown in Fig. 4.10.

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{nk} \quad , k = 0, \dots, N-1 \quad (4.8a)$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)W_N^{-nk} \quad , k = 0, \dots, N-1 \quad (4.8b)$$

$$W_N = e^{-j2\pi/N} \quad (4.8c)$$

where  $W$  is the DFT matrix [159]. A reduction of data points from 301 to 50 is possible without significant visible distortion of the original signal as illustrated in Fig.4.10.



**Figure 4.10: (a) Typical windowed superimposed ECG datasets from the MIT-BIH cardiac arrhythmia database (record no 104, lead V5, 301 sample window), and (b) time domain filtered and reconstructed signal using 50 Fourier coefficients.**

#### 4.4.2.2 Discrete Cosine Transform (DCT)

The DCT is associated with the Fourier transform, which is similar to the DFT using only real numbers. It is based on a transform of an ECG signal into a frequency representation as a sum of components of varying magnitude and frequency. The most common variant of the DCT is the DCT-II type. The DCT-II is typically defined as a real, orthogonal, and linear transformation [160]. A discrete Cosine Transform of  $N$  sample is defined as:

$$C(k) = u(k) \sum_{m=0}^{N-1} x(m) \cos\left(\frac{\pi k(2m+1)}{2N}\right), \quad k = 0, 1, \dots, N-1 \quad (4.9a)$$

where  $x(m)$  the value of  $m^{\text{th}}$  samples of input signals,  $N$  is the number of sample, and  $C(k)$  represents the DCT coefficients.

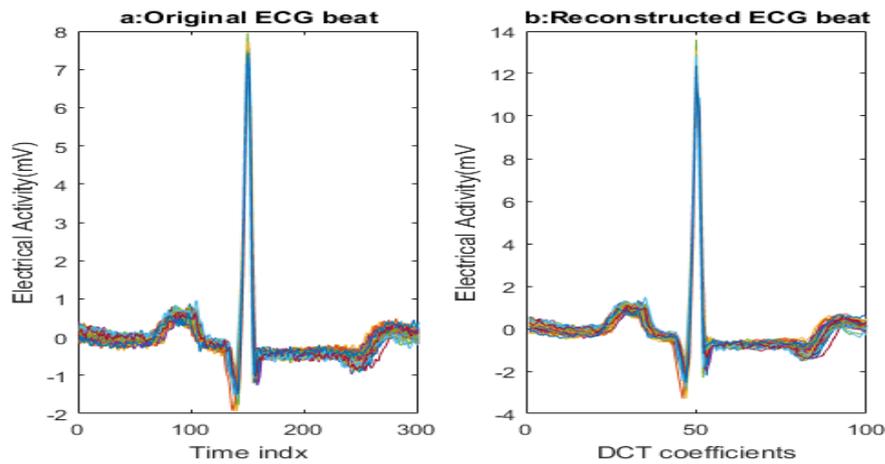
The inverse DCT is defined as:

$$x(m) = u(k) \sum_{k=0}^{N-1} C(k) \cos\left(\frac{\pi k(2m+1)}{2N}\right), \quad m = 0, 1, \dots, N-1 \quad (4.9b)$$

In both Eqs. (4.9a) and (4.9b)  $u(k)$  is defined as:

$$u(k) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } u = 0 \\ \sqrt{\frac{2}{N}} & \text{for } u \neq 0 \end{cases} \quad (4.9c)$$

Fig 4.11 shows an ECG signal taken from the MIT-BIH arrhythmia database and its reconstructed waveform using DCT. One hundred of the original 301 DCT coefficients were used for reconstructing each ECG beat.



**Figure 4.11:** 50 Normal ECG beat extracted from ECG datasets from the MIT-BIH arrhythmia database and b) reconstructed signal using 100 Fourier coefficients.

#### 4.4.2.3 Discrete Sine Transform (DST)

The Discrete sine transform (DST) is one of the Fourier transform methods which is similar to the discrete Fourier transform (DFT) using only a purely real matrix. It corresponds to the complex parts of a DFT of roughly double the length. Discrete sine transforms (DSTs) express a signal in terms of a sum of sinusoids with different frequencies and amplitudes. In addition, a DST runs on a function at a finite number of discrete data points as the discrete Fourier transforms (DFT). The noticeable difference between a DST and a DFT is that the former uses only sine functions in DST, while in DFT both cosines and sines are used [161].

In order to generate the DST, it is required to replace the Neumann condition at  $x = 0$  with the Dirichlet condition. The Dirichlet condition is used to specify the values of unknown function on the boundary of the domain while the Neumann condition specifies the values of the normal derivative on the boundary of the domain. This condition can be used and applied in an exact manner, both at the grid point or at the mid-grid point. These two processes

together with two possible boundary conditions can be applied both to the grid point and at the mid-grid point combined in order to generate the eight DSTs. Even and odd DCTS are derived in the same way. The matrices are denoted as signalled through the DST matrices as  $B_n$ . The size of the grid is considered to be given by  $\pi/N$ . After normalisation by  $\sqrt{2/N}$ , then the unitary matrix elements of the DST can be written as [162].

$$S(k) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \sin\left(\frac{\pi(2n+1)(k+1)}{2N}\right), k = 0, 1, \dots, N \quad (4.10)$$

#### 4.4.2.4 Singular value decomposition (SVD)

Singular value decomposition (SVD) is an matrix transformation in signal processing which is used in spectral analysis, system identification, filter design, applications in statistics, estimation theory, and other related mathematical problems. Scharf (1991) proposed that, given an  $m \times n$  matrix  $A$  of full rank  $r$ , this real ( $m \times n$ ) matrix, where  $m \geq n$  has the decomposition depicted below:

$$A = U \Sigma V^T \quad (4.11)$$

There are three matrices identified as follows: an  $n \times n$  unitary matrix  $V$ , an  $m \times m$  unitary matrix  $U$ , and an  $n \times m$  diagonal matrix with the given diagonal entries which are strictly positive with positive or zero elements, called the singular values.

If the rank of matrix  $A$  is  $r=m=n$  then the matrix of  $\Sigma$  will be reduced to the simple case of  $\Sigma=D$ , and the diagonal components in  $D$  are considered strictly positive ( $d_{ii}>0$ ). Thus, the decomposition is considered effective for both real and complex matrices. The SVD decomposition can be determined by finding the eigenvalue-eigenvector decomposition in terms of the two matrices. In addition, from  $A$ , two positive-definite symmetric matrices,  $(AA^T)$  and  $(A^T A)$ , can be constructed, each of which can be decomposed as follows [163]:

$$(AA^T) = U(\Sigma^2)U^T \quad (4.12a)$$

$$(A^T A) = V(\Sigma^2)V^T \quad (4.12b)$$

The eigenvectors and eigenvalues can be identified using the decomposition above. For example, for  $(A^T A)$ , the eigenvectors are columns of  $V$  eigenvalues and the squared diagonal elements are the eigenvalues.

The SVD of a Matrix  $X$  is the factorisation of this matrix into the product of the three matrices as previously mentioned. In the field of signal processing, SVD has become a key tool for the decomposition of matrices into singular vectors. For example, the ECG beat signal is decomposed with SVD into a singular matrix as  $U$  and  $V$  sub-matrices in which,  $U$  is known as the left singular orthogonal matrix and  $V$  as the right singular orthogonal matrix, while  $\Sigma$  is known as the diagonal singular matrix. These decomposed groups, which contain the singular value of the input matrix of the ECG beat, can be defined using the following equation:

$$X_{N \times M} = U_{N \times N} \times \sum_{N \times M} \times V_{M \times M}^T \quad (4.13)$$

where  $U$  is the left singular orthogonal matrix that contains the amplitude,  $V$  is the right singular orthogonal matrix that contains the basic pattern, matrix  $\Sigma$  is a diagonal singular matrix with the correlation,  $N$  is the length of the beat (sample), and  $M$  is the number of beats.

To retain the desired signal quality, the rank of a singular matrix can be truncated in decreasing order as a low rank matrix ( $b \ll B$ ) where  $B$  is a full rank and  $b$  is the retained rank of a singular matrix. After this truncation, the decomposed singular matrix can be express as follows:

$$X_{N \times M} = U_{N \times b} \times \sum_{b \times b} \times V_{b \times M}^T \quad (4.14)$$

As can be seen, the truncation of rank affects both orthogonal Matrices  $U$  and  $V$  as well as the singular matrix, and the output for each beat is reconstructed with a lower number of coefficients [164].

In addition, SVD is used to perform PCA by decomposing  $X$  to three matrices as shown in Eq. (4.11), and the covariance matrix can be written as follows:

$$Y = \frac{1}{N} U \Sigma^2 U^T \quad (4.15)$$

#### 4.4.2.5 Principal component analysis (PCA)

PCA is one of a linear dimensionality reduction methods which tries to find a data projection direction that shows maximum variability in the dataset [165]. This is a well-established technique that has been used for feature extraction and dimensionality reduction. In 2011, Gupta *et al* indicated that PCA is generally used as a technique for data classification and dimensionality reduction in many a biomedical signal processing applications [166]. In addition, PCA has been applied in digital signal processing and ECG classification [165]. The PCA approach has a number of attractive features: it can reduce the number of dimensions without much loss to the information as shown in Fig 4.12. For this reason, it has been widely used in statistical data analysis, feature extraction, feature reduction, and data compression [10]. The main goal of the PCA is to reduce the dimensionality of the data matrix by finding new variables (PCs). In ECG analysis these are the following main steps in PCA decomposition algorithm [136],[168-170]:

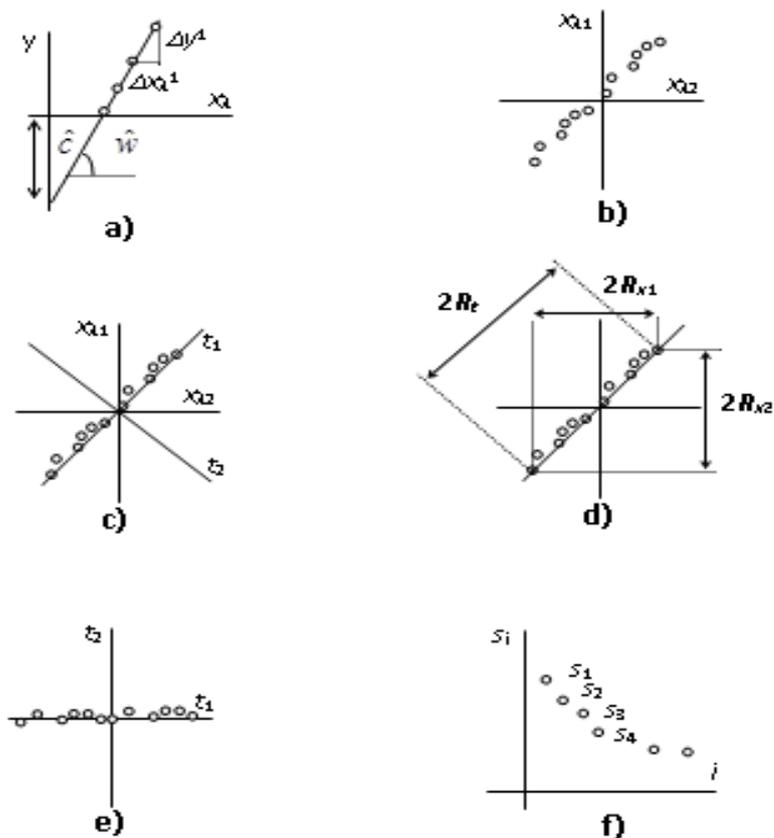
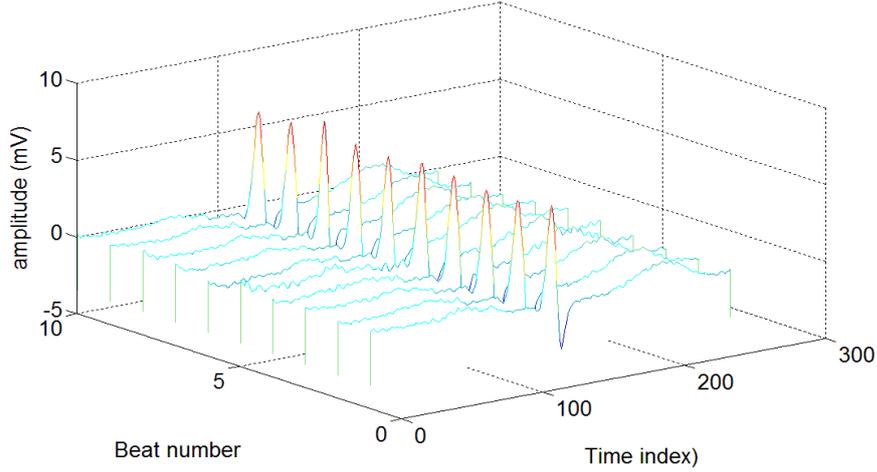


Figure 4.12: a) Evaluation of matrix effect, b) instrument response function at two wavelengths, c) principal component axes, d) data ranges along the original axes and along the first principal component axis, e) rotation along principal component axes, f) reduction of singular components for PC analysis.

- 1- Loading of the data in matrix format after windowing, the data to individual bins (ECG beats): This requires converting a single ECG beat into a data matrix with a column showing the number of beats. Fig 4.13 illustrates the  $N \times M$  data matrix where  $N$  is the number of beats and  $M$  is length of beat segment. In addition, the data is normalized to zero mean and unit variance before applying the PCA algorithm as illustrated in the second phase.



**Figure 4.13: 10 Normal ECG beat extracted from ECG datasets from the MIT-BIH cardiac arrhythmia database record no 209.**

- 2- Calculation of the mean from the data set and subtraction of the mean from the ECG dataset:

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i \quad (4.16)$$

$$Y = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X}) \quad (4.17)$$

where  $Y$  is the variance of a sample,  $X_i$  is the data matrix of a segment of the ECG beat and  $\bar{X}$  is the mean of the dataset,  $N$  is the number of beat samples.

- 3- Calculation of the covariance matrix:

All relationships between pairs of measurements in a data set are described by a covariance matrix [166]. The expression for finding the covariance ( $COV$ ) of  $Y$  is expressed as:

$$COV = \frac{1}{N} \sum_{i=1}^{N-1} (X_i - \bar{X})(X_i - \bar{X})^T \quad (4.18)$$

4- Calculation of the Eigenvectors/Eigenvalues of the covariance matrix:

The following expressions were used to do the sorting of the eigenvalue and eigenvector.

$$COV(S_i) = \lambda_i S_i \quad i \in 1, 2, \dots, p \quad (4.19)$$

with eigenvalues:

$$\det(COV - \lambda_i I) = 0 \quad (4.20)$$

with eigenvectors:

$$(COV - \lambda_i I)S_i = 0 \quad (4.21)$$

where  $S_i$  is the eigenvectors and  $\lambda$  is the eigenvalues of  $COV$  and  $p$  depicts the number of principal components of a given observation vector.

For each eigenvalue  $\lambda$ , the set of all vectors  $\Sigma$  satisfying Eq.(4.19) is called Eigenspace of  $COV$  corresponding to eigenvalue  $\lambda$ . When eigenvectors are found from the covariance matrix, the eigenvalue is used for ordering eigenvectors from highest to lowest value using Eq.4.20. The PCs are obtained by organising the eigenvectors in the descending order of magnitude of eigenvalues. Therefore, only the first eigenvectors are chosen and the final dataset has only dimensions of these eigenvectors that were selected as indicated in the following step.

5- Choosing components and forming a feature vector:

In order to create a *PC feature vector*, some eigenvectors need to be kept from the list of all eigenvectors:

$$PC = \begin{bmatrix} T_1 \\ T_2 \\ \cdot \\ \cdot \\ T_m \end{bmatrix} \quad (4.22)$$

where  $m$  is number of coefficient in each PC.

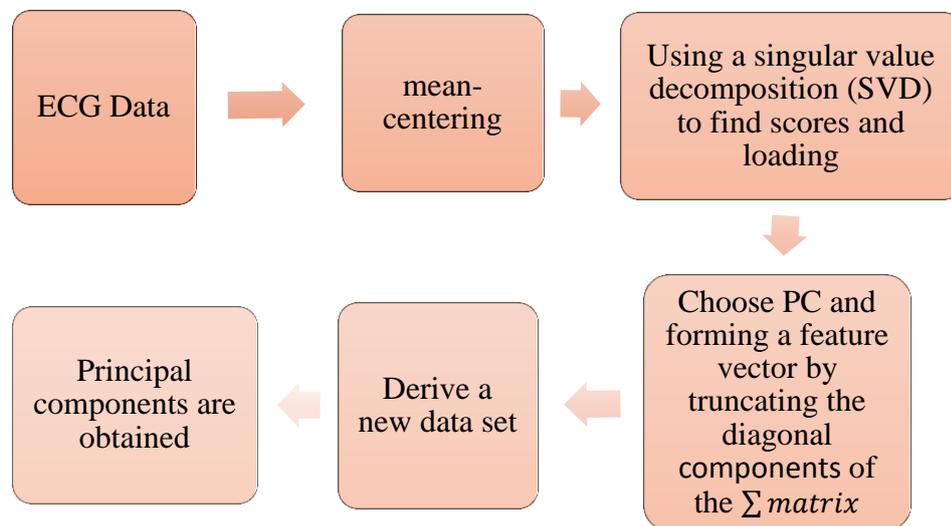
6- Derivation of the transformed dataset the new data set is then performed.

Once the PCs have been chosen and assembled to form the feature vector, the data is fed to the classifier. Projecting the data in the directions of PCs is based on taking the inner product between the data and the eigenvectors. The transpose of the feature vector was obtained in a previous step; this is used to get the final dataset.

$$\text{Final dataset} = \text{PC}^T \cdot DT \quad (4.23)$$

where  $\text{PC}^T$  is the matrix with the eigenvectors placed in the rows due to the transposition operate and  $DT$  is the mean-adjusted data that was obtained in the second step. As can be seen each original data vector can be represented by its principal component vector with reduced dimensionality without missing much information while keeping only important information that is needed for performing classification.

In (2014) Shlens discussed two algebraic solutions for finding the PCs for solving PCA using either an eigenvector decomposition as mentioned above or through a more general solution using singular value decomposition (SVD). In this research both algorithms have been used to select features for ECG signal and reduce dimension of the input vector to the classifier [171]. Fig.4.14. illustrates the main procedure that has been used to calculate the PCA from the ECG signals and create an input to the classifier using SVD.



**Figure 4.14: the main steps of the PCA algorithm using SVD**

The use of PCA is well established in ECG data analysis[114],[172-173]. In 2007 Polat and Gunes proposed using PCA to decrease the ECG arrhythmia beat features from 279 to 15

features per beat. Test classification accuracies of 100% were obtained using these features [174]. In another study [175] different structures including PCA such as PCA-NN and fuzzy c-means clustering (FCM)-PCA-NN were used for classification ECG arrhythmias, where PCA was applied to reduce the sample number in each beat from 200 to 20. Their test results suggested that the FCM-PCA-NN structure showed better result than the PCA-NN structure with an average test error  $5.05 \times 10^{-9}$ .

### 4.4.3 Discussion of the classification algorithm

#### 4.4.3.1 The necessity for a cross-validation step

To minimise the bias that is related to the random sampling of the data samples when comparing the predictive accuracy of two or more methods, one can use a  $k$ -fold cross-validation methodology. This method is also known as rotation estimation; the complete dataset is divided into  $k$  mutually exclusive subsets of about equal size randomly. This way the classification model is trained and tested  $k$  times [176]. Sequentially, each subset is tested using the classifier trained on the remaining  $k-1$  subsets. Thus, each instance of the whole training set is predicted once as a result of the cross-validation accuracy. Using the term accuracy the percentage of data that are correctly classified can be indicated [148].

In stratified  $k$ -fold cross validation, the folds are formed in a way such that they approximately contain the same class labels as the original dataset. The estimation of the overall accuracy of the model of cross validation is measured by simply averaging the  $k$  individual accuracy measures as expressed in Eq. 4.24 [176].

$$ACVA = \frac{1}{k} \sum_{i=1}^k A_i \quad (4.24)$$

where  $ACVA$  stands for cross validation accuracy,  $k$  is the number of folds used, and  $A$  is the accuracy measure of each fold. In the current study cross-validation is used for estimating the best parameter values for the Gaussian kernel function. The parameters with the best cross validation accuracy were selected and used to train the CSVM and SVM.

#### 4.4.3.2 Optimization of SVM parameters and training using Sequential Minimal Optimization (SMO)

SVM training often requires the solution of a complex quadratic programming (QP) optimisation problem. Chunking methods [177], Osuna's algorithm [178] and SMO are all well-established methods that can be used to train SVMs. The following section provides a brief overview of the SMO algorithm that was adopted to train the CSVM to find the hyperplane parameters, threshold value and support vectors in the current work.

The first SMO algorithm was proposed by Platt (1998), this was subsequently enhanced by Keerthi (2001) [179]. It is based on decomposing a complex QP problem into a sequence of smallest-possible QP problems. This, it is a special case of other decomposition methods and it optimises the minimal subset of two samples at each iteration. In traditional SVM, two parameters,  $\alpha_i$  and  $\alpha_j$ , are selected to find optimal values and update the SVM to reflect these new values. Updates are obtained on the basis of the SMO algorithm.

The Karush-Kuhn-Tucker (KKT) conditions can be used to check for convergence of the algorithm to the optimal point. In addition, the optimality conditions for KKT are the gradient of the Lagrange multipliers with respect to the primal variables that must always be satisfied by the dual problem. There are three conditions that depend on the parameter value of  $\alpha_i$  as shown below:

$$\begin{aligned} \text{if } \alpha_i = 0 &\Rightarrow y_i (w^T \cdot x_i + b) \geq 1 \\ \text{if } \alpha_i = C &\Rightarrow y_i (w^T \cdot x_i + b) \leq 1 \\ \text{if } 0 < \alpha_i < C &\Rightarrow y_i (w^T \cdot x_i + b) = 1 \end{aligned} \tag{4.25}$$

where  $(w^T \cdot x_i + b)$  is the output of the SVM for the training samples.

The SMO algorithm provides an efficient method for evaluating the objective function of the dual problem shown in Eq. (4.48a) and for finding the hyperplane parameters. Since SMO belongs to the family of iterative algorithms, optimisation problems are solved by translating this into sub-problems, which are subsequently solved through simpler analytical approaches.

The SMO algorithm provides good performance when several Lagrange multipliers are simultaneously used. The SMOs are considered very efficient methods for training SVMs.

Training SVMs using the SMO algorithm comprises three phases: an analytical method to solve for the two Lagrange multipliers, the use of heuristics to select the two parameters such as  $\alpha_i$  and  $\alpha_j$  for optimisation, and finally a computation of the  $b$  threshold ( $b$  parameter) based on the new parameters.

#### 4.4.3.2.a Application of SMO to evaluate the two Lagrange Multipliers in CSVM

Optimising Lagrange multipliers  $\alpha_i$  and  $\alpha_j$  is done through maximising the objective function of the dual problem and using an old set of feasible solutions, such  $\alpha_i$  and  $\alpha_j$ . First, the SMO computes the constraints on the values of these parameters ( $\alpha_i, \alpha_j$ ) and then solves for the constrained maximisation problem. As there are only two Lagrange multipliers, the constraints can be easily displayed in two dimensions. In Eq. (4.26a) and (4.26b) the bound constraints lead the Lagrange multipliers to lie inside a box-like region, while the linear equality constraint causes the Lagrange multipliers to lie on a diagonal line. The SMO algorithm first computes  $\alpha_j$  and then computes the ends of the diagonal line segment in terms of  $\alpha_j$ . In order for  $\alpha_j$  to satisfy the constraint  $0 \leq \alpha_j \leq C$ , two bounds,  $L$  and  $H$ , must be found, these must satisfy constrains  $L \leq \alpha_j \leq H$ .

If the label value of  $y_i$  does not equal the  $y_j$  value, then Eq. (4.26a) will be applied to  $\alpha_j$ , whereas Eq. (4.26b) will be applied to  $\alpha_j$  if the label value of  $y_i$  is equal to  $y_j$ .

$$L = \max(0, \alpha_j - \alpha_i), \quad H = \min(C, C + \alpha_j - \alpha_i) \quad (4.26a)$$

$$L = \max(0, \alpha_j + \alpha_i - C), \quad H = \min(C, \alpha_j + \alpha_i) \quad (4.26b)$$

In the second step, the  $\alpha_j^{new}$  value must be found to maximise the objective function. The location of the constrained maximum of the objective function is computed. Thus, the value of the first optimal parameter  $\alpha_j$  is computed using the following expression:

$$\alpha_j^{new} = \alpha_j + \frac{y_j(E_i - E_j)}{\eta} \quad (4.27a)$$

$$\begin{aligned} \eta &= K(\bar{x}_i, \bar{x}_i) + K(\bar{x}_j, \bar{x}_j) - 2K(\bar{x}_i, \bar{x}_j) \\ E_r &= f(x_r) - y_r \end{aligned} \quad (4.27b)$$

where  $\alpha_j$  is the old value before optimisation using (4.27a) and (4.27c),  $K$  is a kernel matrix,  $E_r$  is the error between the SVM output on the  $r$ th example and the true label  $y_r$ , and  $\eta$

parameter is the second derivative of the objective function along the diagonal line. Moreover, if the value of  $\alpha_j^{new}$  is outside the bounds of  $L$  and  $H$ , then the value of  $\alpha_j^{new}$  is simply clipped to lie within the boundary range. Hence, the second step clips  $\alpha_j^{new}$  to lie within the boundary range  $[L, H]$  as shown below.

$$\alpha_j^{new,clipped} = \begin{cases} H & \text{if } \alpha_j^{new} \geq H \\ \alpha_j^{new} & \text{if } L < \alpha_j^{new} < H \\ L & \text{if } \alpha_j^{new} \leq L \end{cases} \quad (4.27c)$$

The optimal for  $\alpha_i^{new}$  is given using the  $\alpha_j^{new,clipped}$  as indicated below:

$$\alpha_i^{new} = \alpha_i + y_i y_j (\alpha_j - \alpha_j^{new,clipped}) \quad (4.28)$$

where  $\alpha_i$  is the old value before optimisation.

#### 4.4.3.2.b Selection procedure for SOM parameters

Heuristics are used to select which two Lagrange multipliers are optimised in order to speed-up convergence. There are two separate choices for heuristic selection of the two Lagrange multipliers, (one for each parameter). The outer loop of the SMO algorithm is used as the first heuristic choice and selects the first Lagrange multiplier, whereas the second Lagrange multiplier is selected using an inner loop that maximises the absolute value of the numerator in Eq. (4.27a)  $|E_i - E_j|$ . The outer loop alternates iterations over the entire set of training samples and then selects an example that violates the KKT conditions shown in Eq.(4.25). The outer loop alternates between one sweep through all examples and performs as many sweeps as possible through the non-boundary examples then selects an example that violates the KKT conditions. Since the first  $\alpha_i$  is given, the inner loop looks for a non-boundary value that maximises the step size by the absolute value of the numerator  $|E_i - E_j|$  to select  $\alpha_j$ . For every non-bound sample in the training set, SMO keeps a cached error value of  $E$ , then chooses an error to approximately maximise the step size. The error values are cached because the SMO algorithm spends the majority of its time adjusting the non-boundary

samples. The SMO algorithm selects samples with minimum  $E_j$  if  $E_i$  is positive, whereas if  $E_i$  is negative, it selects a sample with maximum error  $E_j$ .

#### 4.4.3.2.c Computing the threshold that defines the CSVM hyperplanes

The value of threshold ( $b$ ) in both linear and non-linear SVMs must be re-computed after each step and is selected such that the KKT conditions are fulfilled for the  $i^{\text{th}}$  and  $j^{\text{th}}$  examples. After optimising both Lagrange multipliers  $\alpha_i$  and  $\alpha_j$ , the threshold of both  $b_1$  and  $b_2$  are considered valid when the  $\alpha_i^{\text{new}}$  and  $\alpha_j^{\text{new}}$  are not in the bounds ( $0 < \alpha_i, \alpha_j < C$ ). In addition, if both  $\alpha_i^{\text{new}}$  and  $\alpha_j^{\text{new}}$  are not in the bounds after optimisation, the thresholds for  $b_1$  and  $b_2$  can be expressed as shown below:

$$b_1 = E_i + y_i (\alpha_i^{\text{new}} - \alpha_i) K(\bar{x}_i, \bar{x}_i) + y_j (\alpha_j^{\text{new,clipped}} - \alpha_j) K(\bar{x}_i, \bar{x}_j) + b \quad (4.29a)$$

$$b_2 = E_j + y_i (\alpha_i^{\text{new}} - \alpha_i) K(\bar{x}_i, \bar{x}_j) + y_j (\alpha_j^{\text{new,clipped}} - \alpha_j) K(\bar{x}_j, \bar{x}_j) + b \quad (4.29b)$$

However, occasionally, the threshold may be chosen to be the mean value between  $b_1$  and  $b_2$  as illustrated in Eq. (4.29c) when both  $b_1$  and  $b_2$  are valid and equal. Moreover, all the thresholds between  $b_1$  and  $b_2$  satisfy the KKT conditions if both  $\alpha_i^{\text{new}}$  and  $\alpha_j^{\text{new}}$  are at bounds.

$$b = \frac{b_1 + b_2}{2} \quad (4.29c)$$

In the current work, Eq. (4.46a), Eq. (4.47a), Eq. (4.64a), and Eq. (4.65a) form a QP problem that the SMO algorithm will solve individually. The SMO algorithm will terminate when all of the KKT optimality conditions of the QP problem are satisfied.

#### 4.4.3.3 SVM classification

SVM is a reliable classification technique that is based on statistical learning theory. It was developed to classify the dataset that contains two separable classes and is widely used for data classification and function approximation because of its generalization ability. SVM maps input data into a high dimensional feature space where it may become linearly separable. Therefore, dimension reduction is necessary for an efficient classifier, especially for large data

sets. Different approaches for data reduction may be presented for models trained in a supervised way. SVMs are binary classifiers, which consider examples of two distinct classes. In addition, a SVM is a supervised learning technique that uses an optimal separation hyperplane (OSH) to separate classes in binary or multi-class classification problems. SVM is therefore a classifier that works through determining an optimal hyperplane to separates two class data. The optimal hyperplane is the surface optimally separating two groups of data. Moreover, a hyperplane, in an Euclidean plane geometry is a higher-dimensional principle of a line, and in 3D Euclidean geometry it is a higher-dimensional principle of a plane. A hyperplane with the biggest distance to the closest training data point of a functional margin provides a good separation. Generally, when the margin is larger, it means that the classifier's generalization error is lower. The main idea of SVM application is finding the hyperplane that separates two classes in which the margin is maximized between them. The approach is applicable only to linearly-separable training dataset  $x_i$  where  $y_i = \{-1, +1\}$  is a class label.

#### 4.4.3.3.a Binary SVM classification

Binary SVM classification is often performed by using a real-valued function  $f : X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ , where the input  $X = (x_1, \dots, x_n)^T$  is assigned to a positive class if  $f(x) \geq 0$ , and to a negative class if  $f(x) \leq 0$ . The  $f(x)$  function is a linear mapping of  $x \in X$  and can be written as:

$$f(x) = \text{sgn} \langle w, x \rangle + b = \sum_{i=1}^n w_i x_i + b \quad (4.30)$$

where  $(w, b) \in \mathbb{R}^n \times \mathbb{R}$  are the parameters that control the function  $f(x)$ ,  $x_i$  refers to examples (training samples), and  $b$  is the bias term of the hyperplane used to separate binary datasets.

The training set would be:

$$S = ((x_1, y_1), \dots, (x_l, y_l)) \subseteq (X \times Y)^l \quad (4.31)$$

or

$$S = (x_1)(y_1), (x_2)(y_2), \dots, (x_n)(y_n) \mid (x_i)(y_i) \in \mathbb{R}^n \times \{+1, -1\} \quad (4.32)$$

where  $S$  is the training sample,  $l$  is the number of examples and  $x_i$  refers to the *examples* and  $y_i$  are their *labels*. If  $X$  is a vector space, the input vectors are column vectors as are the weight vectors [5]. As there are linearly-separable training and non-linearly-separable training dataset,

two type of SVM need to be considered in binary SVM when Linear performing and non-linear SVM classification.

#### 4.4.3.3.b Linear SVM classification.

Suppose there are linearly-separable training datasets  $x_i$  and  $y_i = \{-1, +1\}$  is a class label. In an SVM formulation wherein the training samples are  $X = \{x_i | i \in \{1, \dots, n\} \wedge x_i \in \mathbb{R}^d\}$  where  $d$  is the dimension of the input vectors and class label  $Y = \{y_i | i \in \{1, \dots, n\} \wedge y_i \in \mathbb{R}\}$  used for training the classifier;  $X$  is linearly separable and  $x_i$  contains two feature components, these examples can be placed in a two dimensional plane. A hyperplane with the upper limit margin for a given finite learning patterns set is called the optimum separation hyperplane (OSH). The OSH problem is defined as:

$$\begin{cases} w \cdot x_i + b \geq +1 & \text{if } y_i = 1 \\ w \cdot x_i + b \leq -1 & \text{if } y_i = -1 \end{cases} \quad (4.33)$$

The above two expressions can be combined into one equation as:

$$y_i (w \cdot x_i + b) - 1 \geq 0 \quad (4.34)$$

Two groups of data can be separated using different hyperplanes. The optimal hyperplane location is right in the middle of the two classes. The two parallel lines in the left and right of the hyperplane are the margins that support the hyperplanes. Fig 4.15 illustrates the structure of simple SVM. The metric  $\frac{b}{\|w\|}$  is the distances from the hyperplane to the basis,  $\|w\|$  is the Euclidean norm of  $w$  and  $d_1$  and  $d_2$  are the distance between the hyperplane and margins.

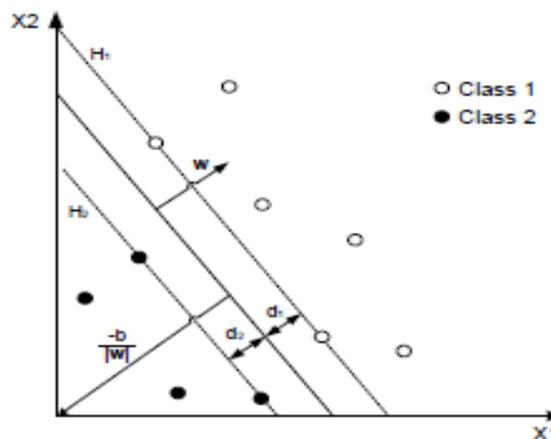


Figure 4.15: Structure of Binary SVM, adopted from [180]

The supporting hyperplanes are defined using Eq.4.34:

The aim in traditional SVM classification problems is to find the optimal separating hyperplane (OSH), which can be obtained by maximizing a margin. In order to maximize a margin, the minimizing of  $\|w\|^2$  need to be found :

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad (4.35a)$$

subject to:

$$y_i (w \cdot x_i + b) - 1 \geq 0, \quad i = 1, \dots, m \quad (4.35b)$$

From Eq. (4.34), it can be found that  $y_i (w \cdot x_i + b) \geq 0$  for  $y_i = +1$ , while  $y_i (w \cdot x_i + b) < 0$  for  $y_i = -1$ .

The optimization problem in Eq. (4.35a) is solved using Quadratic Programming (QP) and Lagrange multipliers. Eq. (4.35a) can be given in the form of a Lagrange functional:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{j=1}^N \alpha_j [y_j (w \cdot x_j + b) - 1] \quad (4.36)$$

Eq. (4.36) provides the basis of the Lagrangian for the optimization problem, where  $\alpha_i$  are the Lagrange multipliers. The goal here is to minimize the Lagrange function Eq. (4.36) with respect to  $w$  and  $b$ , and maximise (4.38) with respect to  $\alpha_i \geq 0$ .

The minimum with respect to  $w$  and  $b$  of the Lagrangian in Eq. (4.36) is given by imposing the following two conditions:

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^N \alpha_i y_i x_i \quad (4.37a)$$

and

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i x_i y_i = 0 \quad (4.37b)$$

The equation below shows the condition satisfying both constrains in (4.37a) and (4.37b), this is also referred within the SVM classifier literature as the dual problem:

$$L(\alpha) = \max \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j x_i x_j \quad (4.38)$$

where  $N$  relates to the number of support vectors;  $L$  is the minimizing goal function in the formulation;  $\alpha_i$  are the Lagrange multiplies;  $x_i$  are the training attribute vectors;  $y_i$  is the training label vectors; and,  $b$  is the bias term of the hyper-plane. The value of  $\alpha_i$  is zero if it is not a support vector and non-zero if it is a support vector. The value of  $\alpha_i$  can be found by solving Eq.(4.38) using quadratic programming (QP) techniques.

The optimal weights of separating hyper-plane is given by:

$$w = \sum_{i=1}^N \alpha_i y_i x_i \quad (4.39)$$

$$b = y_s - \sum_{m \in S} x_m y_m \alpha_m \cdot x_s \quad (4.40)$$

where  $S$  is a set of indices of the support vectors and from each class and  $x_s$  represent support vectors. A hard classifier is obtained using the values of  $w$ ,  $b$  according to the decision function. The optimal decision function (ODF) is express as [181]:

$$f(x) = \text{sgn}(\sum_{i=1}^N \alpha_i y_i x_i \cdot x_s + b) \quad (4.41)$$

If the output of  $f(x) > 0$  is assigned to the positive class, this is considered as a normal ECG beat, whereas if  $f(x) < 0$  it is assigned to the negative class and this is considered as an abnormal ECG beat.

#### 4.4.3.3.c Soft-margin nonlinear support vector machine

Hard margin classification can solve linearly separable data. However, if the data is not linearly separable and imperfect, soft margin classification is used in which some points in the training sample are assumed to be classified incorrectly and a solution is performed without the use of a mapping function. In soft-margin SVM, a slack variable  $\xi_i$  is introduced as shown in Fig 4.16.

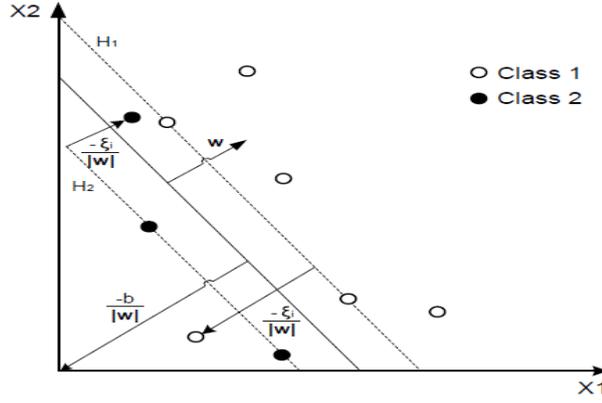


Figure 4.16: Structure of soft margin SVM, adopted from [180]

The OSH hyperplane can be obtained by minimizing the following expression:

$$L(w, \xi) = \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i \quad (4.42)$$

subject to:

$$y_i (x_i \cdot w + b) - 1 + \xi_i \geq 0 \quad \forall i \quad (4.43)$$

where  $C$  is an SVM parameter used for trading off the margin's maximization and classification error's minimization. The Dual formulation for soft margin formulation is represented by the equation written below:

$$L(w, b, \alpha, \xi, \mu) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i [y_i (x_i \cdot w + b) - 1 + \xi_i] - \sum_{i=1}^N \mu_i \xi_i \quad (4.44)$$

Here  $\alpha_i$  and  $\mu_i$  denote Lagrange multipliers. It is necessary to minimize the Lagrangian function  $L$  with respect to  $x$ ,  $w$  and  $b$  and have it maximised with respect to  $\alpha_i$  and  $\mu_i$  to get Eq. (4.44) into a dual problem form. Minimizing Eq. (4.44) is based on differentiating with respect to  $w$ ,  $b$  and  $\xi_i$  in which  $\mu_i \geq 0$  and setting the derivatives to zero. It results to the following three conditions:

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^N \alpha_i y_i x_i \quad (4.45a)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0 \quad (4.45b)$$

and

$$\frac{\partial L}{\partial \xi} = 0 \Rightarrow C = \alpha_i + \mu_i \quad (4.45c)$$

Substituting these condition in Eq. (4.44) to get the Dual formulation leads to:

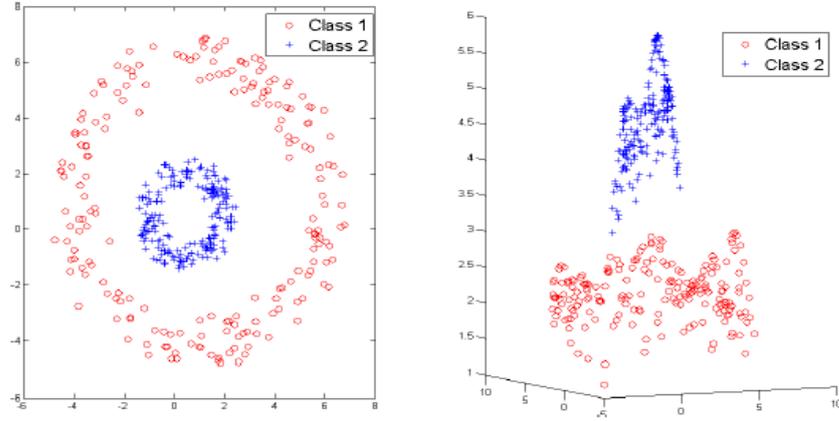
$$L(\alpha) = \max \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j x_i x_j \quad (4.46a)$$

$$\text{subject to: } \begin{cases} \text{one linear equality constraint:} \\ \sum_{i=1}^N \alpha_i y_i = 0 \\ \text{and the inequality constraints:} \\ 0 \leq \alpha_i \leq C \quad i = 1, 2, \dots, N \end{cases} \quad (4.46b)$$

The maximizing and solving of (4.46a) for a given value of  $\alpha_i$  is performed using quadratic programming (QP) techniques. The output is a vector with all Lagrange multipliers evaluated. Each Lagrange multiplier corresponds to a training vector which takes values between 0 and  $C$  ( $0 < \alpha < C$ ), these are called support vectors [181].

#### 4.4.3.3.d Non-linear SVM Classification.

In many applications, using a linear SVM it is not possible to obtain accurate classification results due to the data not being linearly separable. In this case, a suitable mapping function (a kernel function) needs to be used to transform the input data to a higher dimensional feature space where it can be linearly separable as required. Using a kernel function, data in the input space can be transformed to a higher dimensional space. Therefore, in higher dimensional space it is also possible to separate the data using a hyperplane. Fig.4.17 illustrates the difference between an original data set that is not linearly separable in two-dimensional space and a data set that is separable in the nonlinear feature space defined after using a nonlinear kernel function such as a Radial Basis Kernel.



**Figure 4.17: Transformed data to a higher dimensional space and re-mapped using Radial Basis Kernel function, adopted from [180].**

A general non-linear SVM can be expressed as:

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b \quad (4.47)$$

Moreover, for a general non-linear SVM the Lagrangian optimization problem is modified by replacing  $x_i$  with a mapping function  $K(x_i, x_j)$  that performs the non-linear mapping into a feature space as expressed below:

$$L(\alpha) = \max \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (4.48a)$$

$$\text{Subject to} \begin{cases} \text{one linear equality constraint} & \sum_{i=1}^N \alpha_i y_i = 0 \\ \text{and the inequality constraints} & 0 \leq \alpha_i \leq C \end{cases} \quad (4.48b)$$

The decision function of the dual problem for non-linear SVM can be obtained on the basis of the test dataset and the output of Eq. (4.48a) in Eq. (4.49).

A general decision function for non-linear SVM that is used for generating the optimal separating hyperplane in the feature space can be expressed as:

$$f(x) = \text{sgn}\left(\sum_{i=1}^N \alpha_i y_i K(x_i, x) + b\right) \quad (4.49)$$

where  $K(x_i, x_j)$  is a kernel function that measures the similarity of a stored training samples to the input.

#### 4.4.3.3.e Types of Kernel Functions used for non-linearly separable datasets

A kernel method is an algorithm that depends on the data only through inner-products. The kernel function is used to do the evaluation of the inner products in a feature space. In other words, a kernel based solution is composed of a module wherein the mapping is performed into the feature space, and a learning algorithm is subsequently used for discovering the linear patterns in that feature space. This method is well established within the statistical research and machine learning community. The linear patterns can be represented in high-dimensional spaces using the computational shortcut called the *kernel function* [182]. According to Ben-Hur and Weston (2010) [183] the *kernel function* has two advantages. Firstly, it provides the user with the ability to generate non-linear decision boundaries using methods designed for linear classifiers. Secondly, it lets the user apply a classifier that does not have an evident fixed-dimensional vector space representation. Kernel functions can compute the dot product of two-dimensional vectors in an implicitly higher dimensional space and transforms the data from a low dimensional input space to a higher dimensional space [183].

Some of the most commonly used SVM kernels are linear (dot) kernel, polynomial kernel, Gaussian radial basis function kernel, exponential radial basis function kernel, neural kernel, Anova kernel, and Fourier series kernel [184]. In (2010) Albrecht and Darmstadt claimed that the most common kernel functions are: the linear kernel, polynomial kernel, radial basis function (RBF) and sigmoidal kernel [185]. The most common kernels are listed below

##### Linear Kernel

The linear (dot) kernel is represented by the inner product of  $x$  and  $y$  so that:

$$K(x_i, x_j) = x_i \cdot x_j \quad (4.50)$$

This expression is also used for assessing the training set's non-linearity and as a reference for the final classification improvement acquired when using non-linear kernels.

### Polynomial kernel

A polynomial mapping is a general method for non-linear modelling. This is achieved using the expression below:

$$K(x_i, x_j) = \left(1 + (x_i \cdot x_j)\right)^d \quad (4.51)$$

where  $d$  is an adjustable parameter by the user. One disadvantage of this method is that it leads to over fitting as the adoption of a polynomial degree leads to a more complex classification surface.

### Radial Basis Function (RBF) Kernel

The Gaussian Radial Basis Function (RBF) Kernel is of the form:

$$K(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (4.52)$$

Exponential RBF can be used if a lack of coherence in the definition of the hyperplane regions is tolerable.  $\sigma$  refers to the skewness of the distribution and is an adjustable parameter by the user.

### Sigmoidal Kernel:

In artificial neural networks, a hyperbolic tangent function with sigmoid shape can be used. In SVM classification, a sigmoidal kernel can be also used and written using the formulation below:

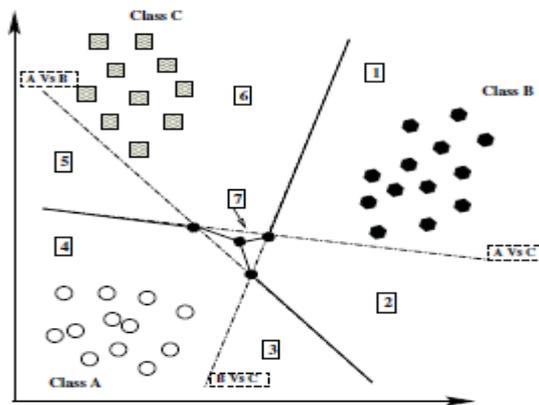
$$K(x_i, x_j) = \tanh\{\gamma(x_i \cdot x_j) + b\} \quad (4.53)$$

#### **4.4.3.3.f Multi-class SVM Classification (MSVM)**

Ultimately, SVMs are designed to solve two-class problems only. However, there are several problems in ECG analysis involving three or more classes. Multi-class SVM problems can be resolved using one of three approaches: one-versus-one (OVO), where all the binary SVMs classify each testing data; one-versus-all (OVA), where the entire formulation is divided into  $k$  binary SVMs; and Directed Acyclic Graph SVM (DAGSVM), which uses a rooted binary directed acyclic graph for prediction, so the class that is less comparable to the input data is

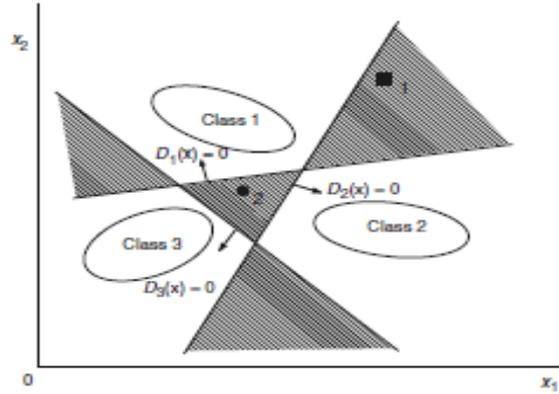
discarded. In subsequent discussion, all formulations of multi-class SVM methods adopt the following general notation:  $x_i \in \mathbb{R}^d$  and  $y_i \in \{1, 2, 3, \dots, k\}$  ( $i = 1, 2, \dots, n$ ), are corresponding class labels [181], where  $d$  is the dimensional training example,  $k$  represents the number of classes, and  $x_i$  are the training attribute vectors.

**One-Versus-One (OVO) Classification:** This technique assumes construction of the normal binary classifiers for all pairs of classes. If  $k$  is the number of classes, then  $k(k - 1)/2$  classifiers are constructed and each one trains data from two classes. All the binary SVMs classify each testing data and the decision function assigns the class that has the largest number of votes. Fig 4.18 provides the structure of OVO to classify three class.



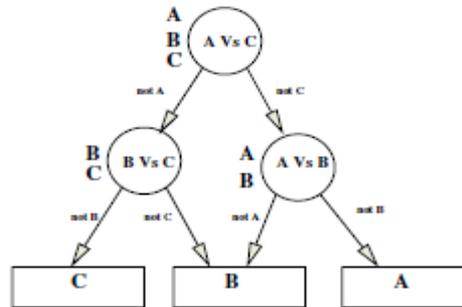
**Figure 4.18: The structure of multi class SVM using of OVO technique, adopted from [186].**

**One versus All (OvA):** This method constructs  $k$  SVM models for a  $k$  class problem. The region is space denoted as 7 in Fig 4.18. Here formulation is divided into  $k$  binary SVMs. All examples in the  $i^{th}$  class that are labelled as positive are used for the  $i^{th}$  SVM trained set and the rest are labelled as negative. The class which has the largest value of the decision function is finally identified. Fig.4.19 provides the structure of OVA to categorize three classes.



**Figure 4.19:** The structure of multi class SVM using of OVA technique, adopted from [187].

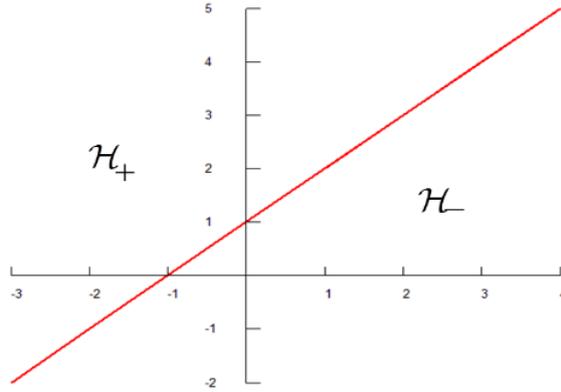
**Directed Acyclic Graph SVM (DAGSVM):** This uses a rooted binary directed acyclic graph for prediction and the class, which is less comparable to the input data, is discarded. This method uses a set of binary SVMs to solve the key problem. The decision function is constructed as a binary tree. Each branch of this tree is a binary SVM associated with for a pair of classes, and in each node of the tree the data is compared with two classes as can be seen in Fig 4.20 [186].



**Figure 4.20:** Structure of DAGSVM algorithm, adopted from [186].

#### 4.4.3.3.g CSVM classification

The purpose of a traditional binary SVM classification task is to separate two different data classes  $\mathcal{H}_+$  and  $\mathcal{H}_-$  using a maximum margin hyper-plane, as shown in Fig.4.21.



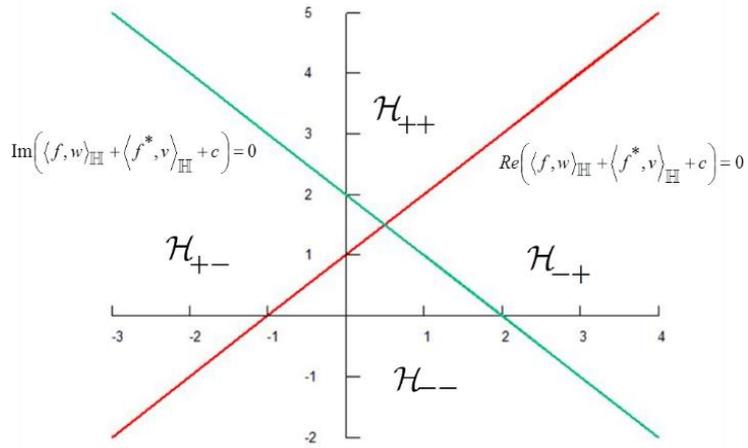
**Figure 4.21: Traditional SVM hyper-plane CSVM.**

In other words, elements belonging to the first class satisfy  $\mathcal{H}_+ = \{f \in \mathcal{H}; \langle f, w \rangle_{\mathcal{H}} + c > 0\}$  true, whereas elements belonging to the second class satisfy  $\mathcal{H}_- = \{f \in \mathcal{H}; \langle f, w \rangle_{\mathcal{H}} + c < 0\}$ . As discussed earlier, in both of the above expressions,  $w \in \mathcal{H}, b \in \mathbb{R}$  are the parameters that control the function and the decision rule is given through Eq.(4.51).

$$\langle f, w \rangle_{\mathcal{H}} + c = 0 \quad (4.51)$$

The adopted formulation for the current study follows the general ideas discussed by Bouboulis *et al.*, (2015) which generalised the SVM formulation to complex spaces. The approach defines a complex space separated into four parts by using a pair of complex hyperplanes. The following section illustrates the CSVM classification technique and procedure that will be used to solve (classify) the four classes' problem in more detail.

For clarity a notation where integers, real numbers, and complex numbers are denoted by  $\mathbb{N}$ ,  $\mathbb{R}$  and  $\mathbb{C}$  respectively is adopted and  $\bar{z}$  denotes the complex conjugate of  $z$ . A complex reproducing kernel Hilbert space (RKHS) will be denoted by  $\mathbb{H}$  while a real RKHS by  $\mathcal{H}$ . The complex hyperplane is composed of two orthogonal hyperplanes that will be referred to as real and imaginary hyperplanes, as shown in Fig 4.22.



**Figure 4.22: complex hyper- plane in CSVM.**

The following expressions provides a complete description of the derived hyper-planes for some  $w \in \mathbb{H}$ ,  $c \in \mathbb{C}$  and  $f \in \mathbb{H}$ :

$$\operatorname{Re}(\langle f, w \rangle_{\mathbb{H}} + c) = 0 \quad (4.52a)$$

$$\operatorname{Im}(\langle f, w \rangle_{\mathbb{H}} + c) = 0 \quad (4.52b)$$

This is directly derived by observing that

$$\begin{aligned} \langle f, w \rangle_{\mathbb{H}} &= \langle f^r, w^r \rangle_{\mathcal{H}} + \langle f^i, w^i \rangle_{\mathcal{H}} \\ &+ i \left( \langle f^i, w^r \rangle_{\mathcal{H}} - \langle f^r, w^i \rangle_{\mathcal{H}} \right) \end{aligned} \quad (4.53)$$

where  $f = f^r + if^i$  and  $w = w^r + iw^i$ .

The above expressions naturally lead to the following conditions associated with the derivation of the hyperplanes:

$$\left\langle \begin{pmatrix} f^r \\ f^i \end{pmatrix}, \begin{pmatrix} w^r \\ w^i \end{pmatrix} \right\rangle_{\mathcal{H}^2} + b^r = 0 \quad (4.54a)$$

$$\left\langle \begin{pmatrix} f^r \\ f^i \end{pmatrix}, \begin{pmatrix} -w^i \\ w^r \end{pmatrix} \right\rangle_{\mathcal{H}^2} + b^i = 0 \quad (4.54b)$$

The above expressions define two separate hyper-planes in  $\mathcal{H}^2$ . These planes are orthogonal if :

$$\begin{pmatrix} -w^i \\ w^r \end{pmatrix}, \begin{pmatrix} -w^i \\ w^r \end{pmatrix} = 0 \quad (4.55)$$

Moreover, in order to be able to define arbitrarily placed hyper-planes intersecting at oblique angles, widely linear estimation functions may also be employed. Through the expressions in Eq.(4.56) below, two hyperplanes associated with doubled real space such as  $\mathcal{H}^2$  may be represented for some  $v \in \mathbb{H}$ ,  $c \in \mathbb{C}$ , and  $f \in \mathbb{H}$  arbitrarily on  $\mathcal{H}^2$  depending on the values of  $w$  and  $v$  by:

$$\text{Re}(\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) = 0 \quad (4.56a)$$

$$\text{Im}(\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) = 0 \quad (4.56b)$$

where  $w, v \in \mathbb{H}$  and  $c \in \mathbb{C}$

A corollary from explicitly adopting complex values leads to alternative expressions for the above equations defining the hyperplanes in a complex space:

$$\left\langle \begin{pmatrix} f^r \\ f^i \end{pmatrix}, \begin{pmatrix} w^r + v^r \\ w^i - v^r \end{pmatrix} \right\rangle_{\mathcal{H}^2} + c^r = 0 \quad (4.57a)$$

$$\left\langle \begin{pmatrix} f^r \\ f^i \end{pmatrix}, \begin{pmatrix} -(w^i + v^i) \\ w^r - v^r \end{pmatrix} \right\rangle_{\mathcal{H}^2} + c^i = 0 \quad (4.57b)$$

Here,  $f = f^r + if^i$ ,  $w = w^r + iw^i$ ,  $v = v^r + iv^i$  and  $c = c^r + ic^i$ .

In CSVM, a complex couple of hyper-planes can be defined as the set of  $f \in \mathbb{H}$ , which can satisfy any of these expression (4.56a) or (4.56b), for some  $w, v \in \mathbb{H}$  and  $c \in \mathbb{C}$ . As can be

seen, the expression (4.52) and (4.56) display the key difference between complex linear estimation and wide linear estimation functions. The complex linear case is pretty limiting, as the couple of complex hyperplanes is always orthogonal, while the widely linear case is broader and covers oblique hyper-plane definitions. In CSVM, a complex pair of hyper-planes, therefore separates the space of complex numbers  $\mathbb{C}$  into four sectors as shown in Fig 4.22 follows:

$$\begin{aligned}
 \mathcal{H}_{++} &= \left\{ f \in \mathcal{H}; \begin{array}{l} \operatorname{Re} (\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) > 0 \\ \operatorname{Im} (\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) > 0 \end{array} \right. \\
 \mathcal{H}_{+-} &= \left\{ f \in \mathcal{H}; \begin{array}{l} \operatorname{Re} (\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) > 0 \\ \operatorname{Im} (\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) < 0 \end{array} \right. \\
 \mathcal{H}_{-+} &= \left\{ f \in \mathcal{H}; \begin{array}{l} \operatorname{Re} (\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) < 0 \\ \operatorname{Im} (\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) > 0 \end{array} \right. \\
 \mathcal{H}_{--} &= \left\{ f \in \mathcal{H}; \begin{array}{l} \operatorname{Re} (\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) < 0 \\ \operatorname{Im} (\langle f, w \rangle_{\mathbb{H}} + \langle f^*, v \rangle_{\mathbb{H}} + c) < 0 \end{array} \right.
 \end{aligned} \tag{4.58}$$

In a CSVM classification task, the primary purpose is the identification of a complex couple of maximum margin hyper-planes that optimally (maximally) separate the points of the four classes, as can be seen in Fig 4.22. Thus, minimisation is necessary using the following equation and maximising margin can be stated by evaluating costing functions through the solution of the following optimisation problem:

$$\begin{aligned}
 & \left\| \begin{pmatrix} w^r + v^r \\ w^i - v^i \end{pmatrix} \right\|_{\mathcal{H}}^2 + \left\| \begin{pmatrix} -(w^i + v^i) \\ w^r - v^r \end{pmatrix} \right\|_{\mathcal{H}}^2 \\
 &= \|w^r + v^r\|_{\mathcal{H}}^2 + \|w^i - v^i\|_{\mathcal{H}}^2 + \|w^i + v^i\|_{\mathcal{H}}^2 + \|w^r - v^r\|_{\mathcal{H}}^2 \\
 &= 2\|w^r\|_{\mathcal{H}}^2 + 2\|w^i\|_{\mathcal{H}}^2 + 2\|v^r\|_{\mathcal{H}}^2 + 2\|v^i\|_{\mathcal{H}}^2 \\
 &= 2(\|w\|_{\mathbb{H}}^2 + \|v\|_{\mathbb{H}}^2)
 \end{aligned} \tag{4.59}$$

Thus, the primal complex SVM optimisation problem is defined as:

$$\min_{w, v, c} \frac{1}{2} \|w\|_{\mathbb{H}}^2 + \frac{1}{2} \|v\|_{\mathbb{H}}^2 + \frac{C}{N} \sum_{n=1}^N (\xi_n^r + \xi_n^i) \tag{4.60a}$$

subject to:

$$\left\{ \begin{array}{l} d_n^r \text{Re} \left( \left\langle \Phi_{\mathbb{C}}(z_n), w \right\rangle_{\mathbb{H}} + \left\langle \Phi_{\mathbb{C}}^*(z_n), v \right\rangle_{\mathbb{H}} + c \right) \geq 1 - \xi_n^r \\ d_n^i \text{Im} \left( \left\langle \Phi_{\mathbb{C}}(z_n), w \right\rangle_{\mathbb{H}} + \left\langle \Phi_{\mathbb{C}}^*(z_n), w \right\rangle_{\mathbb{H}} + c \right) \geq 1 - \xi_n^i \\ \xi_n^r, \xi_n^i \geq 0, \text{ for } n = 1, \dots, N. \end{array} \right. \quad (4.60b)$$

where  $\Phi_{\mathbb{C}}$  is the feature map,  $\mathbf{Z} = x + iy$

The associated Lagrangian function is then casted as follows:

$$\begin{aligned} L(w, v, a, b) = & \frac{1}{2} \|w\|_{\mathbb{H}}^2 + \frac{1}{2} \|v\|_{\mathbb{H}}^2 + \frac{C}{N} \sum_{n=1}^N (\xi_n^r + \xi_n^i) \\ & - \sum_{n=1}^N a_n \left( d_n^r \text{Re} \left( \left\langle \Phi_{\mathbb{C}}(z_n), w \right\rangle_{\mathbb{H}} + \left\langle \Phi_{\mathbb{C}}^*(z_n), v \right\rangle_{\mathbb{H}} + c \right) - 1 - \xi_n^r \right) \\ & - \sum_{n=1}^N b_n \left( d_n^i \text{Im} \left( \left\langle \Phi_{\mathbb{C}}(z_n), w \right\rangle_{\mathbb{H}} + \left\langle \Phi_{\mathbb{C}}^*(z_n), w \right\rangle_{\mathbb{H}} + c \right) - 1 + \xi_n^i \right) \\ & - \sum_{n=1}^N \eta_n \xi_n^r - \sum_{n=1}^N \theta_n \xi_n^i \end{aligned} \quad (4.61)$$

Here,  $a_n, b_n, \eta_n$  and  $\theta_n$  are the positive Lagrange multipliers for  $n = 1, \dots, N$ ,  $N$  is the number of training examples and  $\xi_n^r, \xi_n^i$  are slack variables that permit margin failure. The constraints are given by the saddle point of the Lagrange functional equation and its minimum. The minimum of the Lagrange function in Eq.(4.61) occurs when:

$$\begin{aligned} \frac{\partial L}{\partial w^*} = & \frac{1}{2} w - \frac{1}{2} \sum_{n=1}^N a_n d_n^r \Phi_{\mathbb{C}}(z_n) \\ & + \frac{i}{2} \sum_{n=1}^N b_n d_n^i \Phi_{\mathbb{C}}(z_n) \end{aligned} \quad (4.62a)$$

$$\frac{\partial L}{\partial v^*} = \frac{1}{2} v - \frac{1}{2} \sum_{n=1}^N a_n d_n^r \Phi_{\mathbb{C}}^*(z_n) + \frac{i}{2} \sum_{n=1}^N b_n d_n^i \Phi_{\mathbb{C}}^*(z_n) \quad (4.62b)$$

$$\frac{\partial L}{\partial c^*} = \frac{1}{2} \sum_{n=1}^N a_n d_n^r + \frac{i}{2} \sum_{n=1}^N b_n d_n^i \quad (4.62c)$$

$$\frac{\partial L}{\partial \xi_n^r} = \frac{C}{N} - a_n - \eta_n \quad (4.62d)$$

$$\frac{\partial L}{\partial \xi_n^i} = \frac{C}{N} - b_n - \theta_n \quad (4.62e)$$

when:

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{n=1}^N (a_n d_n^r - i b_n d_n^i) \Phi_{\mathbb{C}}(z_n) \quad (4.63a)$$

$$\frac{\partial L}{\partial v} = 0 \Rightarrow v = \sum_{n=1}^N (a_n d_n^r - i b_n d_n^i) \Phi_{\mathbb{C}}^*(z_n) \quad (4.63b)$$

$$\sum_{n=1}^N a_n d_n^r = \sum_{n=1}^N b_n d_n^i = 0 \quad (4.63c)$$

$$\frac{\partial L}{\partial \xi_n^r} = 0 \Rightarrow a_n + \eta_n = \frac{C}{N} \quad (4.63d)$$

$$\frac{\partial L}{\partial \xi_n^i} = 0 \Rightarrow b_n + \theta_n = \frac{C}{N} \quad (4.63e)$$

Using a Lagrangian, Eq.(4.61) can, therefore, be transformed into its dual form, which is a quadratic QP optimisation problem. The equation below shows the result of using the conditions embodied in Eq. (4.63) with Eq.(4.61), (also known as the objective function of the dual problem). As can be seen, the dual problem in CSVM can be divided into two separate maximisation tasks, as there are two hyperplanes-real and imaginary:

$$\underset{a}{\text{maximize}} \quad \sum_{n=1}^N a_n - \sum_{n,m=1}^N a_n a_m d_n^r d_m^r k_{\mathbb{C}}^r(z_m, z_n) \quad (4.64a)$$

$$\text{subject to} \quad \begin{cases} \sum_{n=1}^N a_n d_n^r = 0 \\ 0 \leq a_n \leq \frac{C}{N} \end{cases} \quad (4.64b)$$

for  $n = 1, \dots, N$

$$\text{and} \quad \underset{b}{\text{maximize}} \quad \sum_{n=1}^N b_n - \sum_{n,m=1}^N b_n b_m d_n^i d_m^i \kappa_{\mathbb{C}}^r(z_m, z_n) \quad (4.65a)$$

subject to

$$\begin{cases} \sum_{n=1}^N b_n d_n^i = 0 \\ 0 \leq b_n \leq \frac{C}{N} \end{cases} \quad (4.65b)$$

for  $n = 1, \dots, N$

Here,  $N$  is equal to the number of support vectors,  $a_n$  is the Lagrange multiplier, and  $z$  the training attribute vector, where  $z = x + iy$ ,  $d_n^r$  is the real training label vector,  $d_n^i$  is the imaginary training label, and  $c$  the bias term of the hyperplane. Furthermore,  $a_n$  and  $b_n$  are Lagrange multipliers, with the value of both zero if it is not a support vector, and non-zero if it is a support vector. The value of both  $a_n$  and  $b_n$  can be found by solving Eq. (4.64a) and (4.65a) using a QP based SMO algorithm [188].

Once the Lagrange multipliers are determined, the real vector  $w$ , imaginary vector  $v$ , and the threshold  $c$  can be derived from the Lagrange multipliers. The hard classifier is obtained using the values of  $w$ ,  $v$ , and  $c$  after placing them into the decision function as shown in (4.66a). In addition, the output of CSVM is explicitly computed from the Lagrange multipliers ( $a_n$  and  $b_n$ ), as can be seen in (4.66b). Eq.(4.66) is a decision function which is used to measure the performance of the classification and find the output of the classifier:

$$g(z) = \frac{\text{sign}}{i} \left( \langle \Phi_{\mathbb{C}}(z), w \rangle_{\mathbb{H}} + \langle \Phi_{\mathbb{C}}^*(z), v \rangle_{\mathbb{H}} + c \right) \quad (4.66a)$$

$$g(z) = \frac{\text{sign}}{i} \left( 2 \sum_{n=1}^N (a_n d_n^r + i b_n d_n^i) \kappa_{\mathbb{C}}^r(z_n, z) + c^r + i c^i \right) \quad (4.66b)$$

$$\text{sign}_i(z) = \text{sign}(\text{Re}(z)) + i \text{sign}(\text{Im}(z)) \quad (4.66c)$$

Kernel functions can compute the dot product of two  $d$ -dimensional vectors in an implicitly higher dimensional space and transforms the data from a low dimensional input space to a higher dimensional space. As discussed earlier, commonly used kernels in the linear kernel,

the polynomial kernel, the Gaussian radial basis function kernel, and the sigmoidal kernel. These can also be adopted for CSVM classification tastes.

In the current study the following general complex Gaussian kernels was used:

$$K_{\sigma, \mathbb{C}^v}(x, y) = \exp\left(-\frac{\sum_{i=1}^v (z_i - w_i^*)^2}{\sigma^2}\right) \quad (4.67)$$

where  $x, y \in \mathbb{R}^v$ ,  $z, w \in \mathbb{C}^v$ ,  $\sigma$  is a free positive parameter chosen by the user, and  $\exp(\cdot)$  is the extended exponential function in the complex domain [189].

#### 4.4.3.4 Soft Independent Modelling of Class Analogy (SIMCA)

SIMCA is a well-established algorithm developed by the chemometrics community, which has gained popularity within the classification community due to the fact that each class model is developed individually. In addition, SIMCA is a “soft” classification process; this implies that a new sample may be classified as a member of one or two or even several classes simultaneously, or even be classified as not belonging to any class. This is in contrast to the “hard” classification process in which a sample is considered classified as a member of one, and only one, class [190].

Very often, in SIMCA the class models are developed using PCA. These models are subsequently reduced to more parsimonious representations though the adoption of a thresholding function which truncates the number of principal components needed to reconstruct the original signals. There are many ways this thresholding value is established, in the current work the predicted residual error sum of squares (PRESS) criterion for each class is adopted, this is evaluated following SIMCA classification and cross-validation.

Creating good models requires an optimization on the number of PCs to be retained. This can be done using a cross-validation procedure wherein the primary model is developed and one PC at a time is added and retained after evaluating the fidelity of the classification result. The residuals for the fit are also calculated at each iteration step. The number of PCs which minimize the residuals is thus optimized so that the best classification result can be obtained. All the different models are optimized in a similar manner for each class represented in the training set. Since each class model is developed independently, these may retain a different number of principal components. After tuning the number of principal components

kept for each class, new samples are used for model validation on the basis of the training set. Scores for each new sample are calculated and projected onto each model. The variance of the residuals for the sample when fit to a class is given from:

$$S_i^2 = \sum_{j=1}^R \frac{(e_{ij})^2}{R-K} \quad (4.68)$$

The following expression is used to compare the variance of the residuals for the sample with the variance of the residuals for members of the class.

$$S_0^2 = \frac{\sum_{i=1}^N \sum_{j=1}^R (e_{ij})^2}{(R-K)(N-K-1)} \quad (4.69)$$

where  $S_i^2$  is the variance in the residuals for the sample fit to the class,  $S_0^2$  is the variance in the residuals for the class,  $e_{ij}$  are the residuals,  $R$  is the dimensionality of the class,  $K$  is the best number the optimal number of PCs retained in the model, and  $N$  is the number of samples in the class [190].

## 4.5 Evaluation of classification results and feature extraction

In classification problems, the primary source of performance measurements output of classification is a confusion matrix. In the confusion matrix, the numbers along the diagonal from upper-left to lower-right denote to the correct classification, and the numbers outside this diagonal represent the misclassification and errors. Overall accuracy is used to denote the overall performance of the classifier. Furthermore, to evaluate the performance of each class in the classification system, some statistical common measures of the performance - sensitivity (SE), specificity (SP), positive predictively (PP) and accuracy (ACC) - are used. The equations of each metrics are given below and these can be calculated using a confusion matrix.

$$\text{Sensitivity (SE\%)} = \frac{TP}{TP + FN} \times 100 \quad (4.70)$$

$$\text{Specificity (SP\%)} = \frac{TN}{TN + FP} \times 100 \quad (4.71)$$

$$\text{PositivePredictivity (PP\%)} = \frac{TP}{TP + FP} \times 100 \quad (4.72)$$

$$Accuracy (ACC\%) = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (4.73)$$

For determining the overall system performance accuracy is usually used as the most crucial metric [176].

$$Overall Accuracy (AC\%) = \frac{Correctly\ classified\ samples}{Total\ number\ of\ samples} \times 100 \quad (4.74)$$

In ECG classification the estimation of the performance of the classifier is based on the recognition of abnormal beats, the true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The *FP* short hand notation refers to a result if a normal class is classified as abnormal, *TP* refers to an abnormal class if it is identified as abnormal, *FN* refers to an abnormal class if it is identified as normal and *TN* refers to a normal class if it is identified as normal. Moreover, the sensitivity (true positive rate) of a test illustrates the percentage of patients in the positive group that are correctly identified, whereas the specificity (true negative rate) states the percentage of non-patients in the negative group that are correctly classified as healthy people.

To evaluate the performance of feature selection the Compression Ratio (CR) and Percentage Root Mean Square Difference (PRD) metrics are used. The CR is defined as the ratio of the number of bits representing the original signal to the number of bits required to store the compressed signal. It is one of the most widely used parameters in data compression algorithms that specifies the amount of compression. A successful algorithm is characterized by a large value of CR. PRD is used to compare between original data and reconstructed. In addition, PRD provides a numerical measure of the residual root mean square error (RMSE) and is used to find out how well the reconstructed waveform matches the original one [191]. In this study, to evaluate the effectiveness of an ECG compression technique CR and PRD were used. The CR is measured by the compression ratio which is defined as the ratio of the number of bits representing the original signal to the number of bits illustrating the compressed signal [170]. The CR and PRD for an ECG signal are obtained using the expression below.

$$PRD = \sqrt{\frac{\sum_{i=1}^n [X(n)_{original} - X(n)_{reconstructive}]^2}{\sum_{i=1}^n [X(n)_{original}]^2}} * 100 \quad (4.75)$$

where  $X(n)_{original}$  is the original ECG signal data and  $X(n)_{reconstructive}$  is the reconstructed ECG signal using PCs.

$$CR = \frac{\text{Total number of samples before compression}}{\text{total number of samples after compression}} \quad (4.76)$$

The techniques described in this chapter are used to generate the result presented and discussed in following chapter.

## 4.6 Summary

This chapter introduced a generic methodology for denoising, feature extraction, and classification of ECG signals. Details on how to remove the high-frequency components of the noise including baseline drift are explained. In addition, several techniques for QRS complex detection are compared. The SIMCA and CSVM algorithm are proposed as suitable candidates for classifying ECG signals with time and frequency domain features respectively. In addition, statistical measures of the performance of the classifier such as sensitivity (SE), specificity (SP), positive predictively (PP), and accuracy (ACC) were discussed and used to evaluate the performance of the proposed algorithms. The following chapter discusses the classification results after adopting the proposed algorithms to perform binary or multi-class classification across a range of known arrhythmia conditions in records found in three different ECG databases.

# **Chapter 5 ECG classification studies: algorithmic implementation, results and discussion**

## **5.1 Introduction**

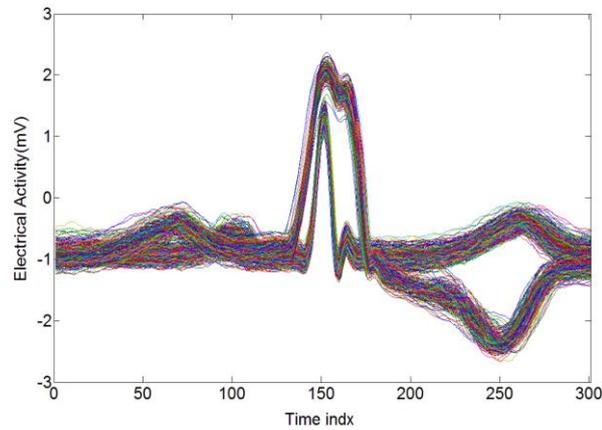
This chapter provides the result of using the proposed CSVM and SIMCA algorithms formulated in the previous chapter to tailor classifier performance. In order to do classification, pre-processing and feature extraction for the input vector that will be presented to the classifier need to be done first. Different algorithms have been used throughout each study in order to extract feature and improve classification accuracy. In some, features were extracted in time domain using PCA, while in other investigations features were extracted in frequency domain using DCT, DST and DFT. In order to extract features in both time and frequency domain Wavelet Transforms (WTs) were used using approximation and details coefficients in time and frequency domain respectively. Finally, an implementation of automatic classification is presented using several algorithms such as the binary and Multi-class Support Vector Machine (MSVM), CSVM and the Soft Independent Modeling of Class Analogy (SIMCA) algorithm.

## **5.2 ECG beat detection and pre-processing**

### **5.2.1 Extraction of normal and abnormal heartbeats**

In this section, the work focuses on individual ECG beat segments (these include P, QRS complex and the T wave). First, the R peak location need to be identified using the accompanying annotation file. Moreover, the associated RR interval can be calculated from the locations of the R points documented in the annotation files of the MIT/BIH database. Moreover, the associated RR interval can be calculated from the locations of the R points documented in the annotation files of the MIT/BIH database. Each beat corresponds to a 0.815 ms segment that includes some signal before and after the R event. The record consists of 301 points at a sampling frequency of 360 Hz. Fig 5.1 illustrates the extraction of normal and abnormal beats with 300 samples around the R peak using MATLAB 2012a. Similar procedures were used to extract beats from the other database. ECGs extracted need to be pre-processed for baseline removal before creating the input vector to be presented to the

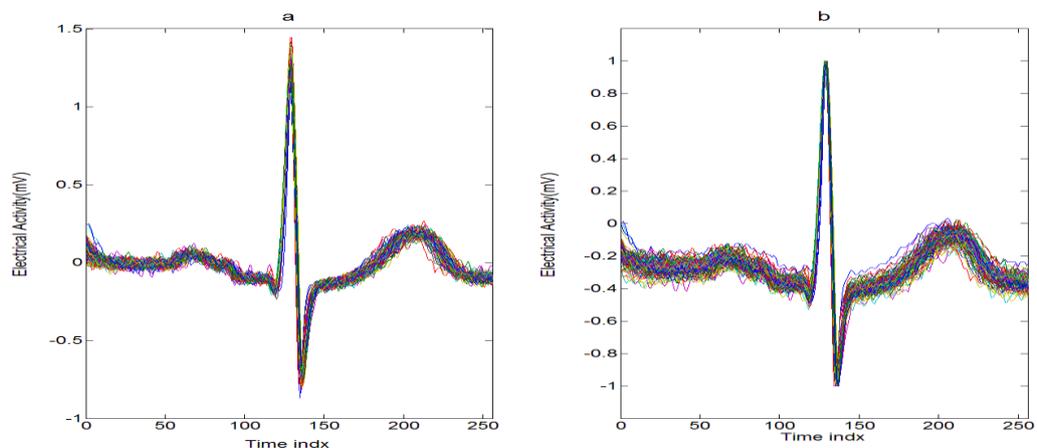
classifier. The Z-score method is utilized to normalize the extracted ECG beat samples with zero mean.



**Figure 5.1: Normal and abnormal beats extracted from the MIT-BIH database using the accompanying annotation files.**

### 5.3 Pre-processing

Pre-processing of ECG beats is performed by subtracting the mean value of each ECG sample in order to eliminate offset effects. This procedure results in normalized signals with zero mean and unity standard deviation as shown in Fig 5.2a. For better classification in order to avoid numerical problems, each beat was normalized between -1 and 1 as shown in Fig 5.2b. The reason for mean centring the data is to decrease possible false decisions due to signal amplitude biases resulting from instrumental or inter-patient variations [9]. In order to improve classification performance the dimensionality of the input vector to the classifier needs to be reduced using feature reduction and selection techniques.



**Figure 5.2: Normal beat extracted from record 100 from the MIT-BIH database with (a) elimination of offset, (b) normalized mean-centred data with removed offset.**

## 5.4 Features extracted and compressed for each beat

Feature extraction and compression are important procedures that usually influence the classification performance of any ECG arrhythmia classification system. Therefore, extraction of sufficient features and a reduction of their dimensions become primary tasks for an ECG arrhythmia classifier in order to achieve optimal classification. The three main ECG feature extraction methods which have been successfully applied for arrhythmia classification are time-domain methods [9], frequency-domain methods [192],[155] and time-frequency domain analysis [94]. All these methods were considered in this project. PCA was used for extracting time domain feature, whereas frequency domain feature were extracted using the Fourier transform. Time-frequency domain features were extracted using DWT as can be successfully applied for arrhythmia feature extraction in both time and frequency domain.

The Principle component analysis (PCA) method has been widely used in statistical data analysis, feature extraction, feature reduction, and data compression in time domain [10]. PCA has been used to find PCs from ECG signals. Complete signal reconstruction of ECG beats is also possible using a few features (PCs) [10]. PCA can be used for feature extraction and dimensionality reduction of the input vector to the classifier.

Several ECG data compression algorithms may be used to reduce the amount of data to be transmitted, stored and analysed, without losing clinical information of interest. The ECG data compression schemes are based on standard transform techniques such as DCT, FFT, DST, and Discrete Wavelet Transform (DWT). According to Deshpande and Rajankar (2013) a compressed ECG signal has advantages such as low storage data space, reduction of low data transmission rate as well as advantages from a transmission bandwidth perspective. Using compression, ECG data can be compressed using a very few number of bits [193]. In (2015) Jain indicated that Data compression can be achieved using the following three approaches: Direct data Compression, Transform Methods, and Parameter Extraction and Compression Methods [170]. Compression approaches can further be divided into the following two main categories: lossless and lossy methods. Lossless methods provide an exact reconstruction of the original signal, whereas lossy methods do not [193]. The second methods has been more commonly used for ECG [194]. ECG compression techniques can be classified as direct time-domain techniques and transformed frequency domain techniques [156]. Transformed frequency domain techniques are based on dividing the signal into frequency components and assigning bits in the frequency domain efficiently. The distribution

of the input signal into blocks of data is done first. The result is then stored as a form of a vector in the frequency domain [157]. Recently several compression methods have been developed with high compression rate and better quality for compressed ECG signal. In (2012) Saberhari and Shamsi investigated ECG signal compression techniques. These included compression in the frequency domain [195]. Yaniv Zigel *et al.* (2000) proposed ECG compression using analysis by synthesis (ASEC) algorithm. This algorithm contains a beat codebook, long and short-term predictors, and an adaptive residual quantizer [196]. Chhipa (2013) investigated a set of ECG data compression schemes for compressing ECG signals. These schemes were based on transform methods such as DCT, FFT, DST, and their improvements. A comparison of performance of these transforms is made in terms of the compression Ratio (CR) and Percent root mean square difference (PRD). It was shown that the DST scheme provided the lowest CR and a high level of distortion, whereas the DCT scheme provided improvement in CR and lowered PRD. However, DCT-II provided an improvement in CR of 95.77%, but it also resulted in an increase in PRD up to 1.33 [197].

In another study [157] five different transform compression techniques DCT-I, DCT-II, DST, FFT, and WT were compared using two criteria: CR and PRD. The simulation results indicated that using a WT with the fourth order Daubechies (db4) function at the second decomposition level provided better CR performance of 97%, and a PRD of 0.8. The study also concluded that the DCT is a valuable tool for compressing both signals in an ECG as well as ECG captured and images because the DCT has de-correlation and energy-saving properties. In addition, the DST is a Fourier-related transform that is similar to the DFT, but it uses a purely real matrix. DST is equal to the imaginary parts of a DFT of roughly twice the length, and it operates on the real data with odd symmetry. Moreover, DSTs can express a function or a signal as a sum of sinusoids with various frequencies and amplitudes [198].

## **5.5 Classification studies using the proposed algorithms**

In the following section, beat classification is performed by using binary SVM and MSVM using OAO (One against One) training, CSVM and SIMCA. Each study will be discussed in details. Each method in this section has been already discussed in detail in chapter 4.

### 5.5.1 Study one: using binary SVM classification results

The ECG beat dataset used to conduct binary SVM classification described below were obtained from the MIT-BIH Arrhythmia Database. An automatic classification of normal and premature ventricular contraction (PVC) beats, is implemented. Fig 5.3 illustrates in a block diagram form the binary classification scheme that was used to detect normal from abnormal ECG beats. Each cardiac cycle or beat in the ECG is normally characterized by a sequence of deflections that make waveforms that are known as the P wave, the QRS complex, and the T wave. QRS detection is performed with the WFDB Software Package to read annotation and find the R (peak) location. Principle component analysis (PCA) and wavelet decomposition using different mother wavelets is used for features extraction and compression. Between 10 and 30 principal components were selected for each beat in the feature selection phase. A different number of ECG beats from two leads and patients are selected for training and testing implementation. The binary SVM experiments were presented at the 8<sup>th</sup> Saudi Students Conference in the UK 2015 as a poster.

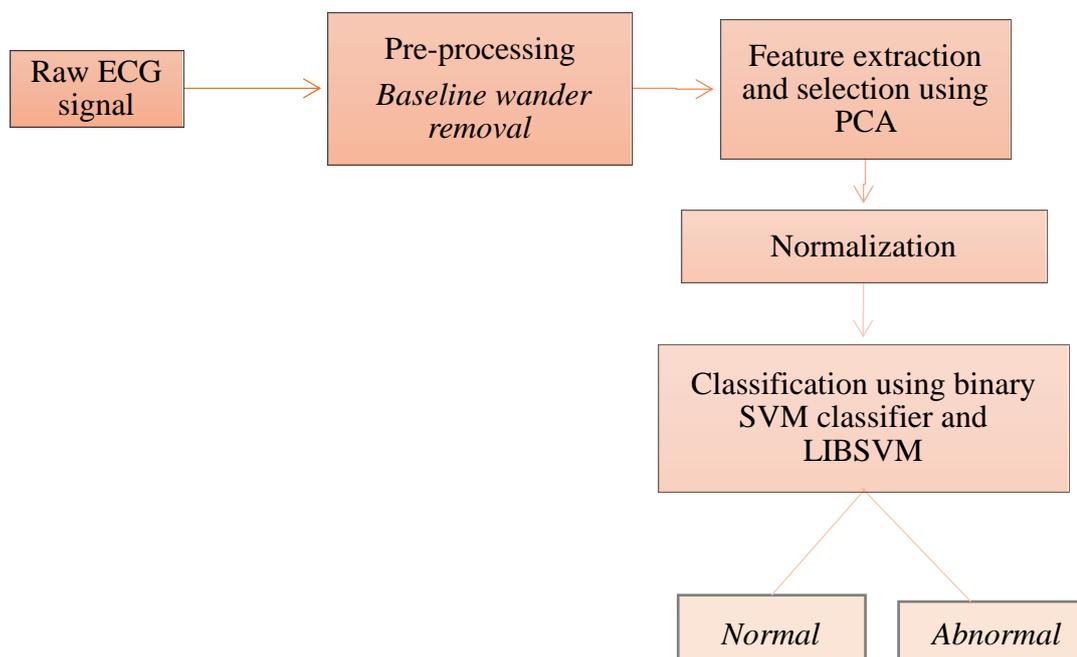
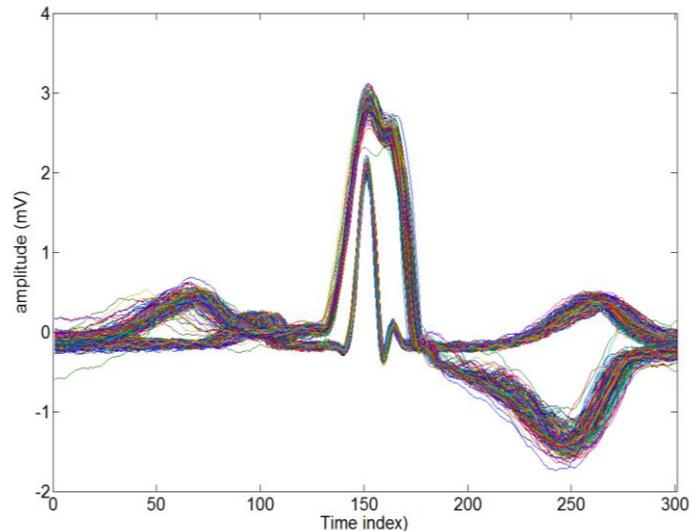


Figure 5.3: Block diagram of binary ECG LIBSVM classification.

### 5.5.1.1 Results using the MIT-BIH Arrhythmia Database in datasets from one lead

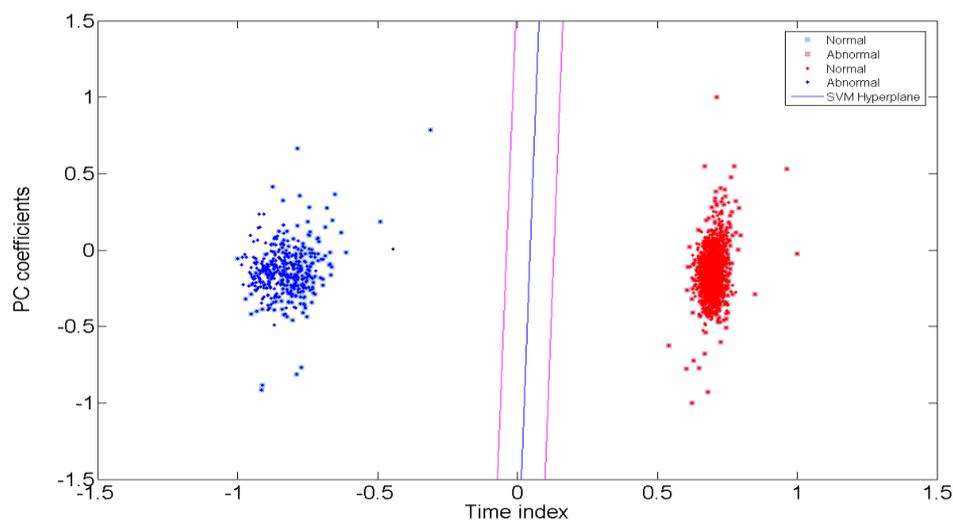
In this study the ECG signals were taken from Lead II (MLII). First, the position of the R wave in the QRS complex was detected using the WFDB Software Package which is normally used for reading annotation of beats and finding the R-peak location. After finding the position of the R-peak in the QRS complex, ECG beats are extracted using windows with 150 samples before the R-peak and 150 points after the R-peak of the QRS complex. The R peak corresponded to the 151<sup>st</sup> point in the data sequence as shown in Fig 5.4. The total length of the segment was approximately 830 ms. This figure illustrates 300 Normal and 200 PVC beats extracted. The normal and PVC beats that were extracted corresponded to one lead. They corresponded to patient record number 119. These beats were selected for both training and testing.



**Figure 5.4:** 300 Normal and 200 PVC beats were extracted from one lead and patient rec.119. The time index is 301 points long corresponding to a time duration of 830ms.

PCA is another useful statistical technique for finding patterns in datasets of high dimensionality. PCA can be used to reduce the dimensionality of the dataset without losing much information [170]. In this study, extracting features from each beat in the time domain was also performed using PCA. Between 10 and 30 principal components (PCs) were selected from each beat. These PCs form the input feature vector to the classifier. The optimum number of principal component was found as 20. The obtained training patterns, whose size was dimension included 20 features times 500 segments, were presented to the binary SVM classifier.

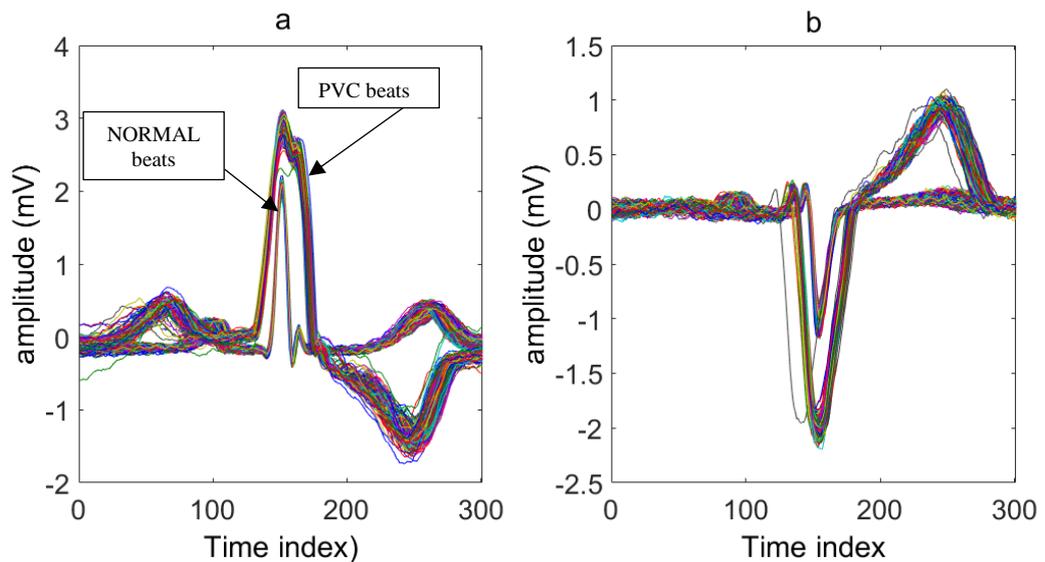
Classification performance was done using a binary SVM algorithm and the LIBSVM software. Figure 4.1 illustrated the general procedure used for implementation through LIBSVM and MATLAB 2012a. A typical LIBSVM software contains two main stages: training a data set to obtain a model and using the model to predict information of a testing data set. In more detail, the procedure consists of the following steps: using LIBSVM to transform data to the format of an SVM package; conducting simple scaling on the data; choosing one of the kernel function and cross-validation to find the best parameter value; using the best parameter value to train the whole training set; and finally creating the model and using a second group (test data) for final evaluation. The main benefit of scaling is to avoid attributes in larger numeric ranges dominating those in smaller numeric ranges (data conditioning). This is achieved by linearly scaling each value within the range between -1 and +1. Cross-validation (in our case 5-fold cross-validation), is used for finding the best parameter value for the kernel function (such as Gaussian RBK). The parameters with the best cross validation accuracy were selected as best. After the best parameters are found, these values are applied to the entire training set to create the model. Implementation of SVM is performed by dividing the dataset into two groups: a train dataset and a test dataset. A training SVM model was created using the training dataset, label, best parameters, kernel function and binary SVM implementation. Finally, the model and test data were adopted for classifying a new ECG beat (test dataset). Fig 5.5 shows the classifier result using 20 PCs as features and binary support vector machines classification with accuracy 100%. The blue squares illustrate normal beats, whereas the red squares illustrate abnormal beats in the data test.



**Figure 5.5: LIBSVM classification result from beats extracted from one lead.**

### 5.5.1.2 Results using the MIT-BIH Arrhythmia Database in datasets from two leads

In this experiment the ECG beats were selected from one patient's record containing two different leads limb (II) and precordial (V1) leads. Pre-processing of the ECG signals included QRS complex and beat detection in a similar manner. One thousand heart beats (600 normal and 400 abnormal) were extracted from MIT-BIH Arrhythmia Database record number 119 lead (II) and lead (V1). Five hundred beats were taken from each lead. Fig 5.6 shows the normal and abnormal ECG beats extracted from two leads after pre-processing. For feature selection from individual beat segments, PCA was used and each beat was reconstructed using only 20 PCs instead of 301 samples. After feature selection, the dataset was divided into two groups: train dataset and test dataset. Five hundred beats were used for training and Five hundred for testing. In order to classify the signals between normal and abnormal classes, binary SVM was used.



**Figure 5.6:** ECG beats extract of beats from two leads (a) limb lead (II) and (b) precordial lead (V1).

This classifier was applied to the reduced space containing only the ECG PC features. During the training phase, the binary SVM parameters are selected using a 5-fold cross-validation procedure. In addition, 5-fold cross-validation was used to obtain the optimal Gaussian RBK kernel parameters. These parameter values were then used to the whole training set to create the SVM model. In order to evaluate the performance of the classification, the SVM model was then adopted for the test dataset samples. Using the binary SVM classifier and PCA for discriminating normal from abnormal beats showed an excellent performances with accuracy ratio 100%. Fig 5.7 displays the classifier result using ECG beats

extracted from the two leads shown in Fig 5.6 on the basis of 20 features for each beat. Green and black squares show the result of normal and abnormal beats when these were taken from precordial lead V1 respectively, while red and blue square illustrate the classification result of normal and abnormal beats when these were taken from limb lead (II) respectively. As can be seen, the values of normal and abnormal beats were significantly different between leads; thus, it is better to make separate models for each lead when creating a dataset for training and validation. Furthermore, the characteristics of beats can be different between patients, as shown in the next study.

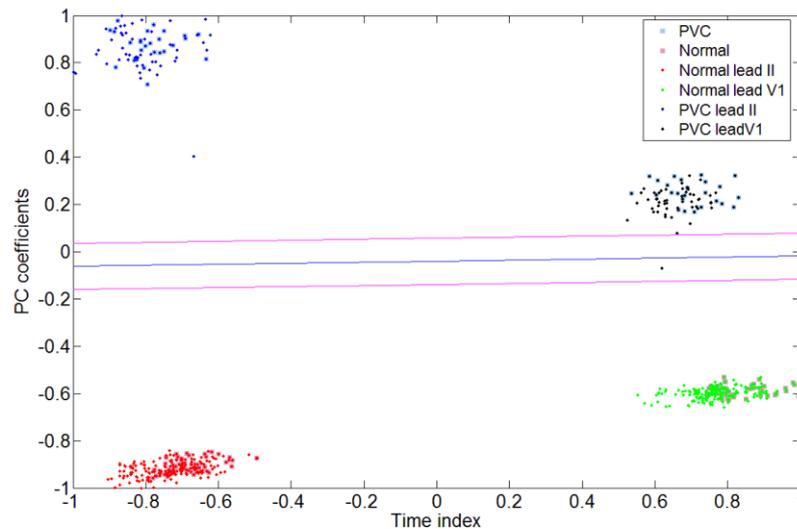
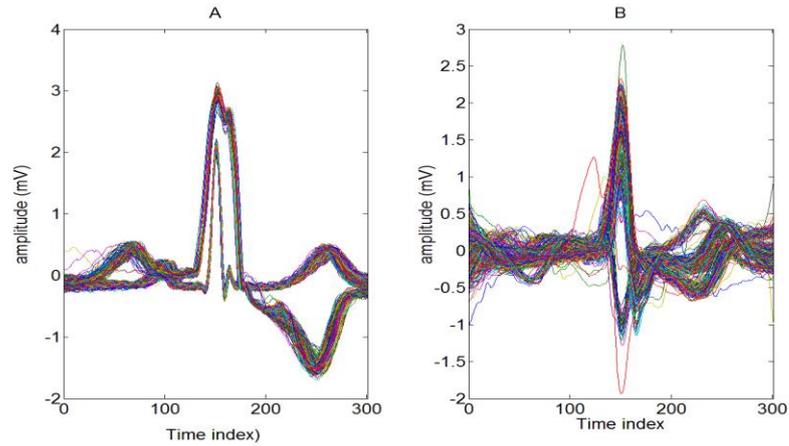


Figure 5.7: Classification result were beats extracted from two leads.

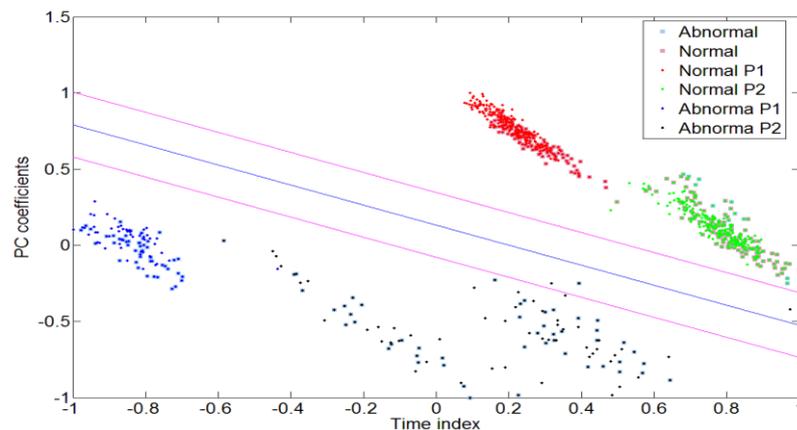
### 5.5.1.3 Results using MIT-BIH Arrhythmia Database in datasets from two patients

In this study, one lead ECG beats were extracted from two different patients. Pre-processing was performed using a similar method to that in the previous study. One thousand ECG heartbeats (500 normal and 500 abnormal) were selected from the MIT-BIH Arrhythmia Database (lead (II) record number 119 and record number 223). Five hundred ECG beats (normal and abnormal) were extracted from each patient\_record. Fig 5.8 show the normal and abnormal ECG beats extracted from two patient records, these can appear dramatically different.



**Figure 5.8:** ECG beats extract from two patients and limb lead (II) (A) patient record no 119 and (B) Patient record no. 223.

The features were extracted in the time domain using PCA and 20 PCs were used for reconstructing individual beats to be presented to the input to the classifier. After ECG beat detection and feature selection, the datasets were divided into two groups: training dataset and test dataset. The training dataset was used for training the SVM and creating the model. The model and test data were combined in order to predict the status of a new ECG beat. Fig 5.9 illustrated the classifier result on the basis of 20 features and binary support vector machines classification which showed an accuracy of 100%. The blue squares illustrate normal beats, whereas the red abnormal beats. As can be seen the characteristics of normal and abnormal beat were different between patients; thus it is better to create a dataset for both training and testing using an extraction procedure from both different patients as well as different leads. A further investigation of this argument can be seen in the multi-class study were a number of beats were extracted from different leads and patients the evaluation of the algorithm.



**Figure 5.9:** Classification result for two patient's information.

#### **5.5.1.4 Discussion of SVM binary classification**

In the current study, a binary SVM classification approach was developed for the classifying normal and abnormal ECGs beat with PCA for feature extracting. This proposed method included (filtering, feature extraction and classification) which are done automatically.

According to current research, several binary classification methods have been developed in the last few years in an attempt to simplify the task of ECG monitoring. Cascade forward NNs is one of the neural networks that can be used to identify whether an ECG beat is normal or abnormal. These types of neural network are similar to feed-forward networks, but they include a weight connection from the input layer to each following layer, and then from each layer to the successive layers. A good illustration of this is a three-layer network which has connections from layer 1 to layer 2, layer 2 to layer 3, and layer 1 to layer 3. Ayub and Saini (2011) identified normal from abnormal ECG beat using a cascade forward network algorithm. The result of classification indicated that using a cascade-forward network with back propagation training is the best option for identifying normal beats, they showed that this type of NN achieved an accuracy of about 99.9%. A further advantage of relevance to mobile application is that the memory requirements are also low. In addition, in that work, the authors discussed these results in relation to previous studies that used other neural network techniques, and concluded that this method gives the best results [89].

Artificial neural networks (ANN) and SVM have been used to diagnose an ECG signal as either healthy or as indicating myocardial infarction (MI). Back propagation artificial neural networks with varying hidden layers and nodes have been implemented for performance analysis. The Pan-Tompkins algorithm has been used to detect the QRS complex, while feature extraction and the reduction of the dimensionality of the input vector to the classifier have been performed using DWT and PCA. Using SVM provided a better classification than when using NN, with an overall accuracy of 91.0714 % and 90.1786% for approximate and detail coefficients, respectively. The neural network was shown to have an overall accuracy of 82.14% for the approximate coefficient and 78.1% for the detail coefficient [199].

SVM was used to classify the ECG signals into normal and arrhythmia categories with an accuracy of 94%. Features such as linear predictive coefficients (LPC), linear predictive cepstral coefficients (LPCC) and mel-frequency cepstral coefficients (MFCC) were extracted

to present the ECG signal to the classifier. Using SVM with MFCC features provided better performance results with an accuracy rate of 94%, while other features, such as the predictive coefficients (LPC) and the linear predictive cepstral coefficients (LPCC), showed a classification accuracy rate of between 90 and 93% [200]

In 2006 Exarchos *et al* investigated an alternative methodology based on ‘association rules’ for the automated detection to classify ECG beats as either ischemic or normal. Electrocardiogram (ECG) features were extracted from the ST segment and the T-wave. Simulation result showed 87% sensitivity (SE) and 93% specificity (SP) respectively [201].

Kanaan *et al.* (2011) proposed a combination of support vector machines and the principal component to classify normal and abnormal beats into two classes. The principal component analysis was used to reduce the dimension of the ECG beat, while the kernel principal component analysis was used for high-dimensional mapping of nonlinear separable data. The binary SVM classifier was used to classify normal and abnormal classes. Using a binary SVM classifier with KPCA as a feature extraction method gave better performance results than did PCA, with accuracy, sensitivity, positive predictively and specificity of 95%, 100%, 90% and 90%, respectively [169].

A neural network with a back propagation algorithm was used to classify arrhythmia conditions according to normal and abnormal classes with an accuracy rate of 96.77% on the MIT-BIH database [202]. In another study, a supervised neural network was designed to distinguish between normal and ischemic beats of the same patient. It was based on an ECG digital recording that was taken from the European ST-T database and two beat segments (RRR interval). Experimental results showed a highly reliable beat classification and the automatic detection of ischemic episodes with a specificity of 99% and sensitivity of 98% [203]. In [4], a classical multilayer feed-forward neural network with a back propagation algorithm was used to classify normal and arrhythmic beats with an accuracy rate of 100%. This study showed successful classification results that agree with the current study’s classification results.

A nonlinear principal component analysis (NLPCA) algorithm and multilayer neural network technique was developed and used to extract nonlinear features and to classify segments of beats as normal or abnormal. A specific part of the ECG beat called the ST segment was selected and used with only two nonlinear features for each ST segment as

inputs for the classifier. Using the NLPCA showed a better performance than did the linear principal component analysis (PCA), with a classification accuracy rate of 80% for the normal beats and 90% for the ischemic beats [114].

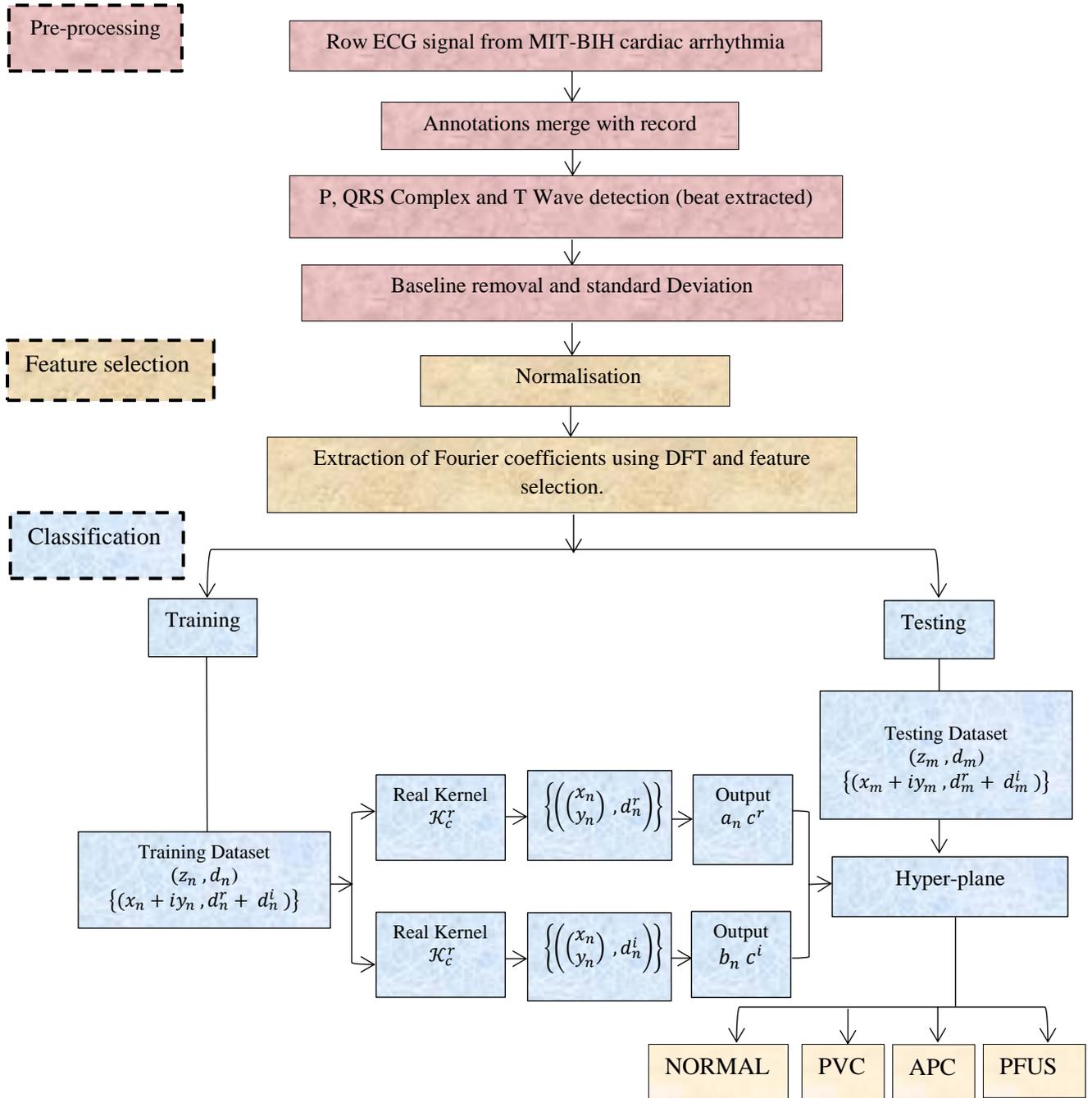
This study showed that using a binary SVM with PCA for feature extraction improved the accuracy of normal and abnormal beat classification. It is clear that using binary SVM with a kernel function operates as an excellent classifier for the given normal and abnormal arrhythmia beat data set. Compared to some previous studies, the proposed methodology produced better results than did other approaches in terms of accuracy.

### **5.5.2 Study two: classification using CSVM and MSVM classifiers**

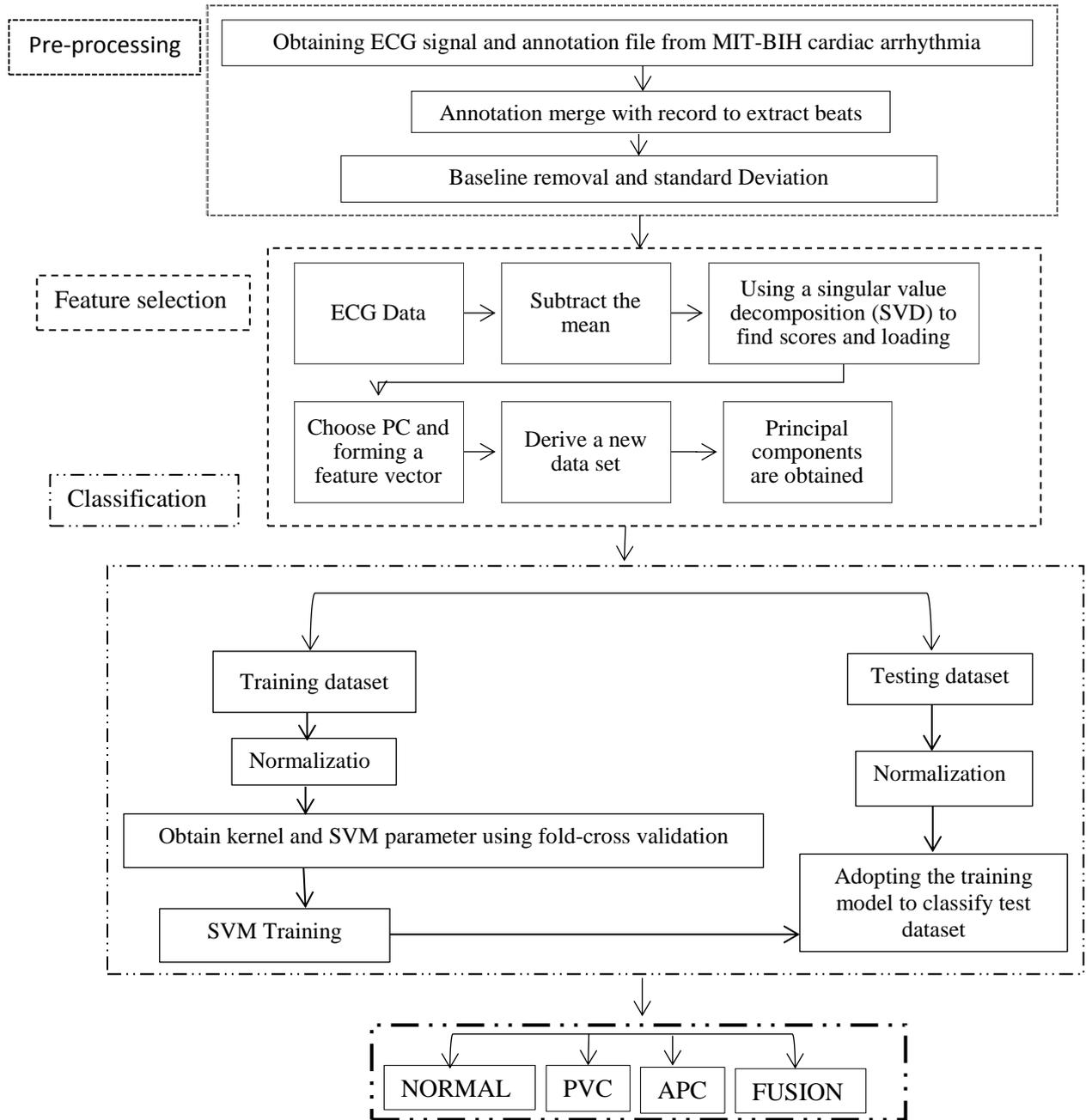
The aim of this study has been to improve multi-class SVM accuracy by extending traditional SVM algorithms to complex spaces so as to simultaneously classify four types of heartbeats. Different ECG beats, normal and abnormal were obtain from the two most popular ECG databases (the MIT-BIH arrhythmia database and European ST-T database). A description of the method, and a table showing a confusion matrix with results of the performance of classification is provided. A discussion of how the results contribute to the overall study and improve how the classification accuracy is then articulated. The proposed arrhythmia classification scheme is composed of the same pre-processing steps, with beat extraction, feature reduction and selection and finally a classification step as shown in the block diagram in Fig 5.10. Fig 5.10 and 5.11 illustrate description of the three most important steps associated with CSVM and MSVM classification respectively.

#### **5.5.2.1 Results on MIT-BIH Arrhythmia Database using a precordial lead**

The 301 sample window function used in the ECG analysis corresponds to a 360-Hz sampling rate, furthermore this is again centered around the R peak of a single ECG beat sample. The extracted ECG beat samples are normalized using Z-scores and standard deviation is used to normalize the signal amplitude. The MIT-BIH Arrhythmia database contains forty-eight half-hour records that were obtained from twenty-five males and twenty-two females. Each record was sampled at eleven-bit resolution and had a duration of thirty minutes [51].



**Figure 5.10: Overview of the pre-processing, feature selection and classification steps associated with the proposed CSVM algorithm used in the multiclass ECG beat classification problem.**



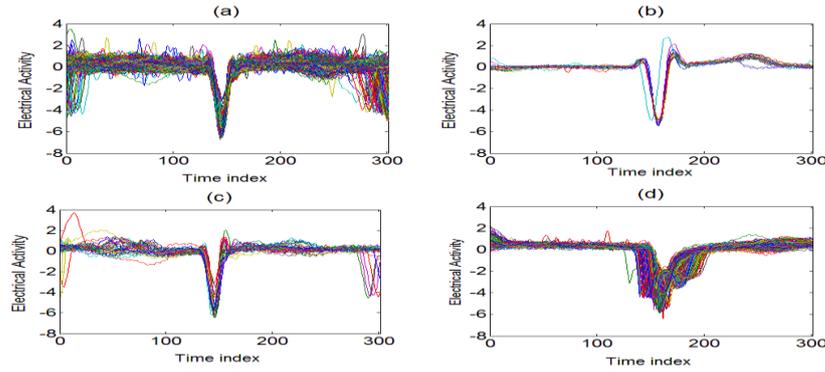
**Figure 5.11: Overview of the pre-processing, feature selection and classification steps associated with the proposed SVM algorithm.**

For the current study, twelve recordings and four ECG beat types were only considered. Table 5.1 illustrates details of four dataset groups that are used in this experiment in terms of record number, lead name and total number of beat samples for each beat type studied.

Dataset	Annotation	Record #										
			One lead	Two lead		three lead			four lead			
			V1	V1	V5	V1	V5	V2	V1	V5	V2	V4
Train	NORMAL	201, 202, 210, 217, 223 100, 102, 104, 114, 123 102, 104, 117	100	50	50	35	33	32	35	33	32	0
	PVC	201, 202, 210, 217, 223 102, 104, 114, 123 102, 104, 124	50	25	25	24	23	3	16	16	3	15
	APC	201, 202, 223 100, 114, 117, 124	50	30	20	30	20	1	30	19	1	1
	PFUS	217, 102, 104	50	25	25	18	16	16	18	16	16	0
Test	NORMAL	201, 202, 210, 217, 223 100, 102, 104, 114, 123 102, 104, 117	100	50	50	35	33	32	35	33	32	0
	PVC	201, 202, 210, 217, 223 102, 104, 114, 123 102, 104, 124	50	25	25	24	23	3	16	16	3	15
	APC	201, 202, 223 100, 114, 124	50	29	20	29	20	0	29	19	0	1
	PFUS	217, 102, 104	50	25	25	18	16	16	18	16	16	0
Total		12	500	250	250	213	184	103	197	168	103	32

**Table 5.1: Illustration four-dataset groups extracted from MIT-BIH arrhythmia database.**

At this point, it is worth mentioning that various studies have been performed in the past in order to determine appropriate sizes of the training and testing datasets of the ECG signals. Übeyli, (2007) showed beat classification accuracies of 98.61% on the basis of 360 training beats and 360 test beats. Similarly in the current study, half of the beat samples are selected as the training dataset and the other half is used for testing in each group. All pathological ECG records such as PVC, APC, and PFUS associated with different leads were contrasted with normal beats. Each cardiac cycle in an ECG signal consists of the P-QRS-T waves. Detection of the QRS complex plays a central role in ECG classification. Annotation files were used to assign the position of the R waves as it was determined in some previous study [204],[5] so that the QRS complex can be easily detected. The WFDB (Waveform Database) software package with library functions (from PhysioToolkit) is used for the evaluation of digitized signals with annotations. QRS complex waves as well as T and P waves are identified using both file annotation information of the R peaks and record information. The R-peak location in the QRS complex were also detected using the WFDB Software Package. Following identification of the R-peak, ECG beats were extracted after inclusion of 150 samples at 415 ms before the R-peak with a window spanning 150 data points at 415 ms after the R-peak as shown in the Fig 5.12 similarly to previous studies [89]. Subsequently, the extracted ECG beat signals are normalized using the Z-score method. All raw ECG beat signals used in this experiment were obtained from precordial leads.



**Figure 5.12:** Extract of (a) Normal, (b) PVC, (c) APC and (d) PFUS beats.

The ECG features were extracted in the frequency domain as complex datasets need to be used with CSVM. The features of the P, QRS complex, and T waves are associated with the location, duration, amplitudes, and shapes of the waves. As stated in chapter 3, there are several dimensionality reduction and frequency domain feature extraction algorithms that may be applied successfully to ECG signals such as wavelet transforms (WT) [205] and Fourier transforms [122]. In the current study, the DFT is used to select features in each beat. Before performing a DFT transformation, the ECG signals were normalised; this is in contrast to traditional MSVM classification studies where ECG beat normalization is performed after feature selection [147]. Linear scaling is performed to avoid collinearity issues when reconstructing individual ECG beats in the feature selection phase.

All signal processing was carried out using MATLAB 2012a and WFDB software packages. For classification and CSVM implementation, the reconstructed individual ECG beats with Fourier coefficients were further divided into two groups for training and testing purposes. The optimal hyper-plane parameters are identified on the basis of the training data set, class label and the output of the SMO algorithm. Finally, on the basis of the training results, test data are imported to the CSVM classifier to perform unknown beat classification. The CSVM classifier is used to distinguish the four ECG arrhythmia types using the DFT coefficients as feature vectors. The Gaussian Radial Basis Function parameters are tuned using a cross-validation technique. In addition, CSVM parameter  $C$  and Gaussian Radial Basis Function parameter  $\sigma$  are selected according to a 5-fold cross-validation procedure that was used to find the optimal parameters. Once the best parameter values are obtained, these are then used on the entire training data set. Through trial and error, it was concluded that the best parameter for the kernel function and CSVM module are 170 and 13 for  $C$  and  $\sigma$  respectively. These values provided a cross-validation accuracy rate of 98.40%.

The SMO algorithm was used to train the CSVM and solve the dual problem associated with Eq. (4.64a) and (4.65a). The output of the CSVM consisted of  $a_n$  and  $b_n$  Lagrange multipliers associated with the real and complex parameters in the calculated hyperplanes respectively; the corresponding threshold values for both real and imaginary hyperplanes can be seen in Fig. 5.10. The outputs from training ( $a_n$ ,  $b_n$  and threshold) were combined with the test dataset to measure the classification performance attained and find the output of the classifier.

To evaluate the performance of the CSVM classifier, records from one, two, three and four leads were grouped and presented at the input of the classifier simultaneously. The classification performance was evaluated after presenting 500 beats as associated with the four ECG beat types. There were 200 Normal, 100 PVC, 100 APC and 100 PFUS beats; these were associated to different patient records as shown in Table 5.1. Patients records were randomly divided into two equally sized groups consisting of 250 beats for training and 250 beats for testing. To evaluate the performance of the classification process, three common measures- SE, SB, and PP are used. Overall accuracy is used to denote the overall performance of the classifier.

All four dataset groups used in this study have been trained with the same value of  $C$  and  $\sigma$ . In order to systematically assess the best number of features needed to be selected from the DFT coefficients, different numbers of coefficients were used as input vectors and the classification process was repeated. Table 5.2 compares the error rate associated with different numbers of DFT coefficients as the input vector to the CSVM classifier. It can be seen that using 50 DFT coefficients produced the best result with 0.8% error rate, whereas using 301 DFT coefficients produced the worst result with a 3.5% error rate.

DFT coefficients number	Error rate
301	3.50
100	1.75
50	0.8
30	1.75

**Table 5.2: Classification error rate for different number of DFT coefficients presented to classifier**

The classification results for each ECG beat type are displayed using confusion matrices, which illustrate the success of the classification process in terms of classified or misclassified results; these are shown in Tables 5.3, 5.4, 5.5 and 5.6. In addition, it is worth noting that any ECG beats that were associated to PVC, APC and PFUS are classified correctly, and there is no misclassification in all the dataset groups. The performance of the CSVM using a real Gaussian kernel, in terms of specificity, sensitivity and Positive Productivity on the test sets is illustrated in Table 5.7. This shows the percentage correct classification of individual classes (beat types) in terms of SE, PP and SP for each class. As can be seen in the four groups (datasets), the PP and SP classification results achieved for the APC beats using fifty DFT coefficient were the same 100%, whereas the PP and SP results achieved for NORMAL, PVC and PFUS beat classification results were not as high. For beats extracted from a single lead using the DFT filtered beats and CSVM classifier with DFT features, NORMAL beats could be classified with a specificity of 98%, whereas PVC, APC and PFUS could be classified with a specificity of 100%. Furthermore, using ECG beats from two leads can reduce the number of misclassifications of NORMAL beats. According to Table 5.7, the average sensitivity for NORMAL beats extracted from one, four and three leads has improved from 98%, 96% and 98% respectively to 100% when beats were extracted from two leads. Finally, as can be seen in the results in Table 5.7, the CSVM algorithm can also classify simultaneously the four ECG beats more effectively. Moreover, the overall accuracy using ECG beats extracted from one, two, three and four leads were 99%, 100%, 98% and 99% respectively.

Annotation	Output result			
	NORMAL	PVC	APC	PFUS
NORMAL	98	0	0	2
PVC	0	50	0	0
APC	0	0	50	0
PFUS	0	0	0	50

**Table 5.3** classification results using DFT extracted coefficients of the ECG signal taken from one lead and CSVM.

Annotation	Output result			
	NORMAL	PVC	APC	PFUS
NORMAL	96	4	0	0
PVC	0	50	0	0
APC	0	0	50	0
PFUS	0	0	0	50

**Table 5.5** classification results using DFT extracted coefficients of the ECG signal taken from three lead and CSVM.

Annotation	Output result			
	NORMAL	PVC	APC	PFUS
NORMAL	100	0	0	0
PVC	0	50	0	0
APC	0	0	50	0
PFUS	0	0	0	50

**Table 5.4** classification results using DFT extracted coefficients of the ECG signal taken from two lead and CSVM.

Annotation	Output result			
	NORMAL	PVC	APC	PFUS
NORMAL	98	2	0	0
PVC	0	50	0	0
APC	0	0	50	0
PFUS	0	0	0	50

**Table 5.6** classification results using DFT extracted coefficients of the ECG signal taken from four lead and CSVM.

Annotation	Classification performance result											
	One lead			Two lead			three lead			four lead		
	<i>ST</i> (%)	<i>SP</i> (%)	<i>PP</i> (%)	<i>ST</i> (%)	<i>SP</i> (%)	<i>PP</i> (%)	<i>ST</i> (%)	<i>SP</i> (%)	<i>PP</i> (%)	<i>ST</i> (%)	<i>SP</i> (%)	<i>PP</i> (%)
NORMAL	98	100	100	100	100	100	96	100	100	98	100	100
PVC	100	100	100	100	100	100	100	98	93	100	99	96
APC	100	100	100	100	100	100	100	100	100	100	100	100
PFUS	100	99	96	100	100	100	100	100	100	100	100	100

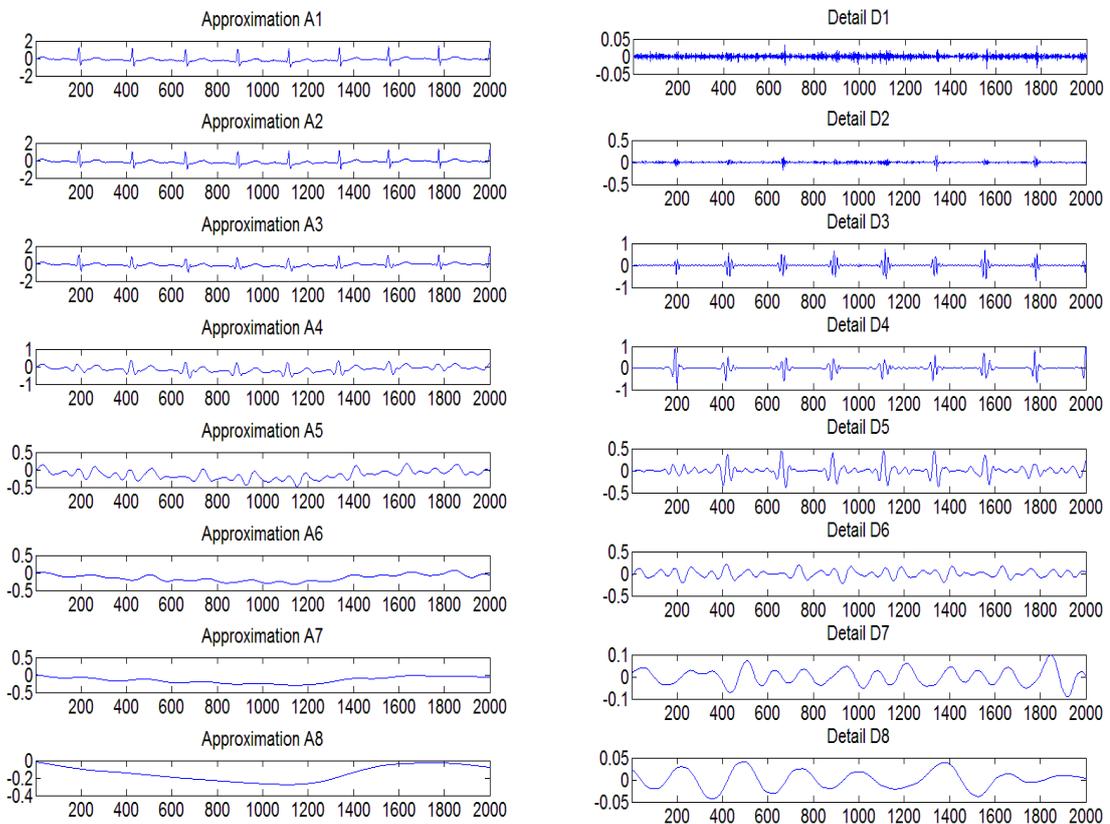
**Table 5.7** Collective performance analysis and classification result using DFT and CSVM.

### 5.5.2.2 Classification of NORMAL, PVC, APC and FUSION beats from the MIT-BIH Arrhythmia Database (using a single limb lead).

In this study, a new automatic classification algorithm to simultaneously classify four ECG beat types, NORMAL, PVC, APC and FUSION is implemented using MSVM and CSVM classification [26]. The ECG signals used in these studies were obtained from the MIT-BIH cardiac arrhythmia database using limb lead (II) from two patient records (record 208 and record 209). The MSVM and CSVM classifier are used to distinguish the four ECG arrhythmias types using the DCT, DST and DFT coefficients as feature vectors at the input vector of the classifier. A sequential minimal optimization (SMO) approach is used to train the CSVM and compute the corresponding complex hyper-plane parameters. These procedures were selected for training and evaluation of classifier performance.

In pre-processing, wave detection (P, QRS complex, T) is performed using Multiresolution Wavelet Analysis [206-207],[151]. One of the advantages of using wavelet decomposition is that it helps in the removal of noise and baseline wander. The multiresolution wavelet for ECG wave detection and feature extraction is performed using DWT and Daubechies 6 (db6) wavelet as Daubechies wavelets have physical similarity with the ECG wave, especially the QRS complex. In addition, in low frequencies their energy spectra are more focused. The morphological features of the ECG signal such as, PR interval, PT interval, ST interval, TT interval, QT interval, and P, Q, R, S, T and peaks points that were detected and used to create the segmentation of the beats. First, the DWT is applied to the ECG record. Decomposition of the signal is done up to level eight as indicated in [208]. Fig 5.13 shows the reconstruction of detail and approximation coefficients at each level of decomposition. QRS complex detection is performed first by using detail coefficients d3, d4, and d5 to detect the R peak, and then Q and S points are identified using more detail

coefficients from other levels. Fig 5.14 illustrated the reconstruction information from d3, d4 and d5 detail coefficients. After the R peak is detected, the Q and S points are detected using d2, d3, d4 and d5 coefficients. The reconstructed wave involving these coefficients is shown in Fig 5.16. Fig. 5.15 and 5.17 show the detected R peak and Q point from the QRS complex. This approach obviates the need to use annotation files. The ECG beats obtained had 301 samples. The reconstruction coefficients d6 and d7 are used to detect T and P waves. The reconstructed wave from both d6 and d7 is shown in Fig 5.18. T and P wave detection were illustrated in Fig 5.19 and Fig 5.20 respectively. Fig.5.21 provides four types of beat extracted for testing and evaluation of the classification algorithm.



**Figure 5. 13: Reconstruction of detail and approximation coefficients at each level of decomposition.**

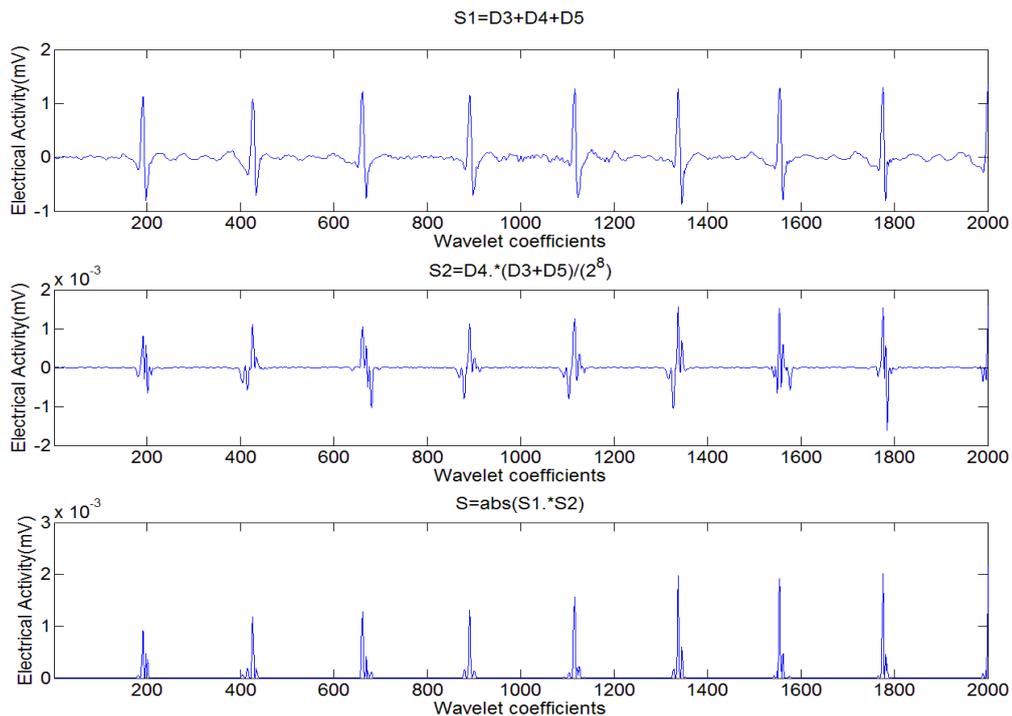


Figure 5.14: Reconstruction using d3, d4 and d5 details coefficients of heartbeat.

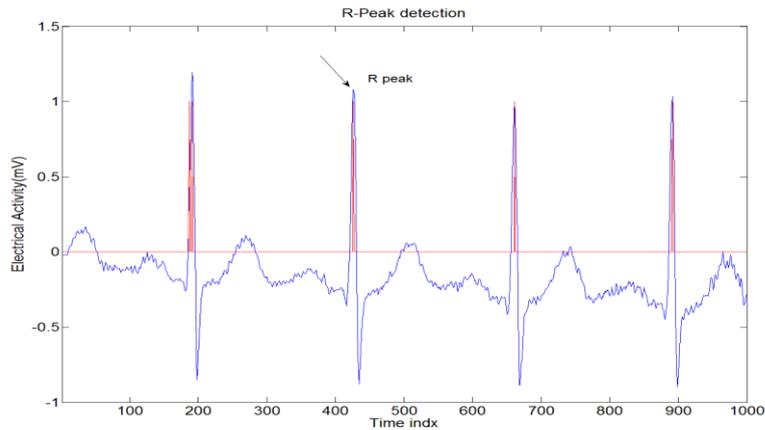
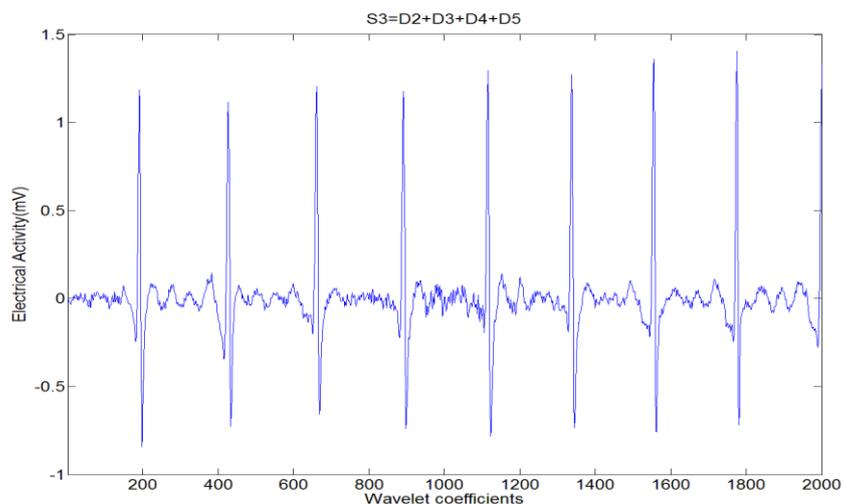
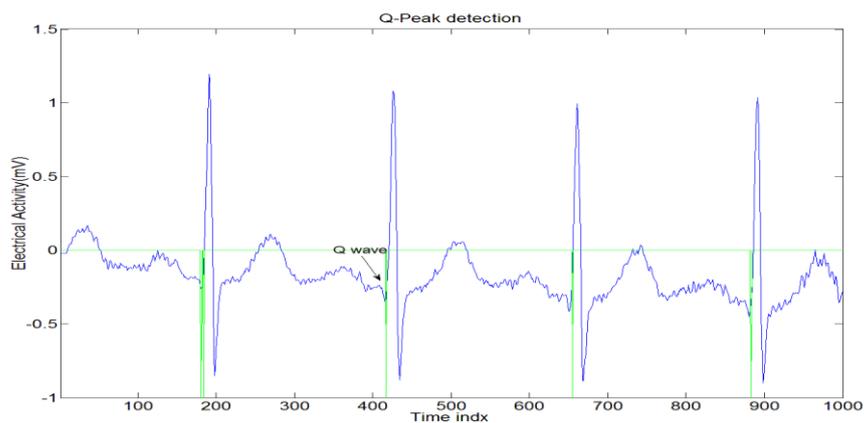


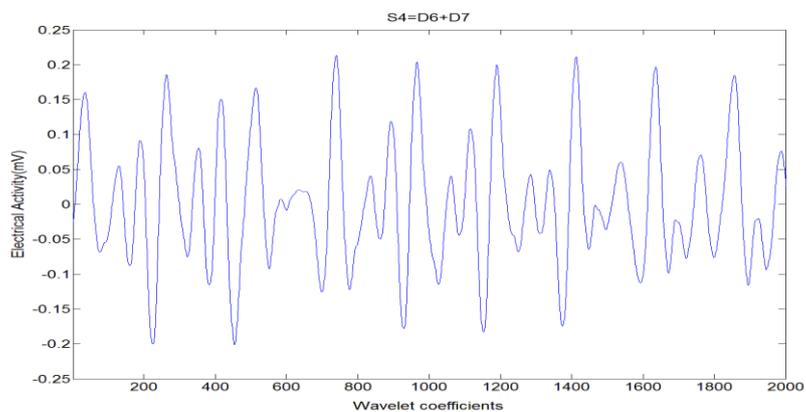
Figure 5.15: Detection of the R peak.



**Figure: 5.16 Reconstruction of the heartbeat using of d2, d3, d4 and d5 details coefficients.**



**Figure 5.17: Detection of the Q peak.**



**Figure 5.18: Reconstruction wave using d6 and d6 details coefficients.**

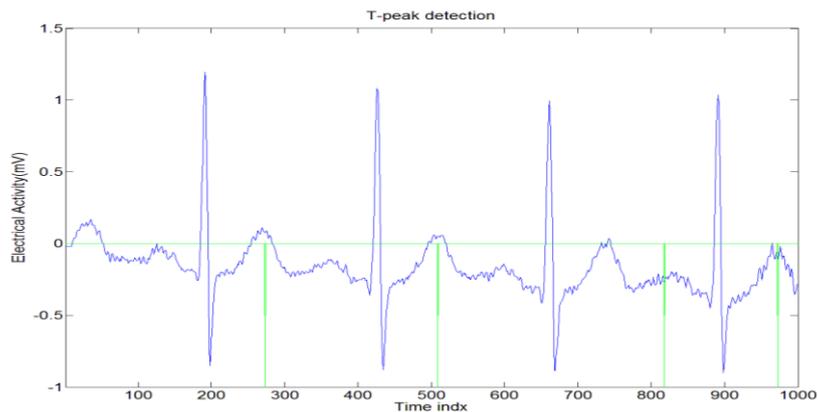


Figure 5.19: Detection of the T wave peak.

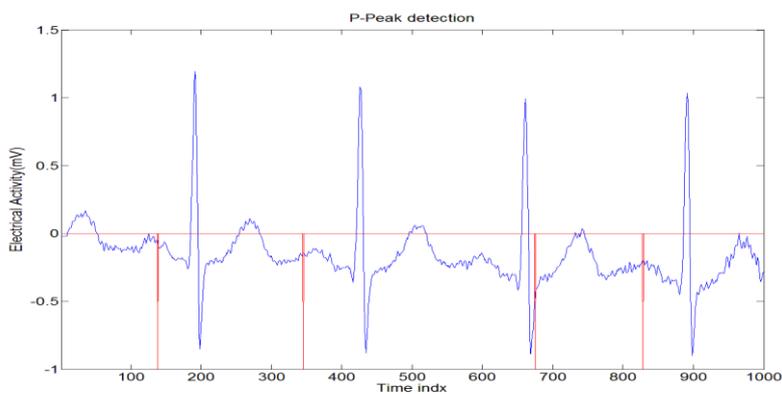


Figure 5. 21: Detection of the P peak.

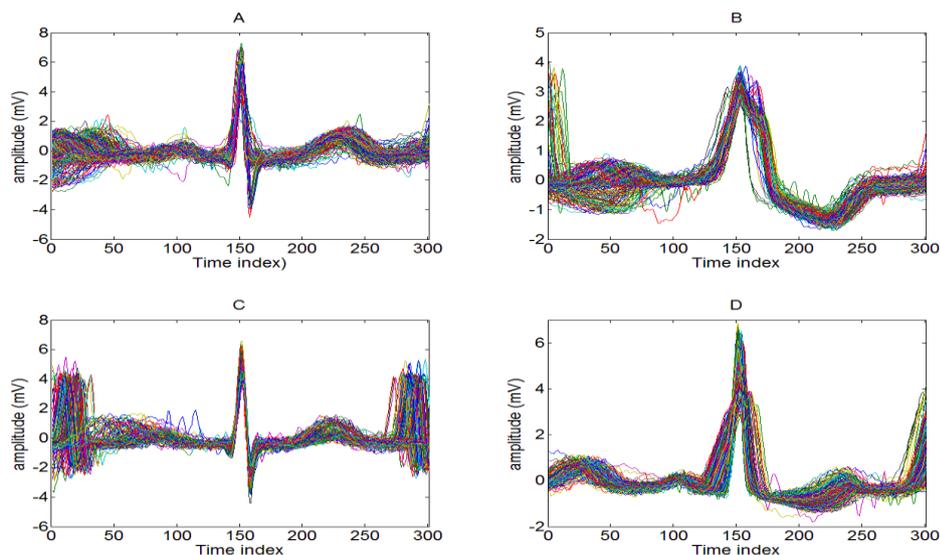
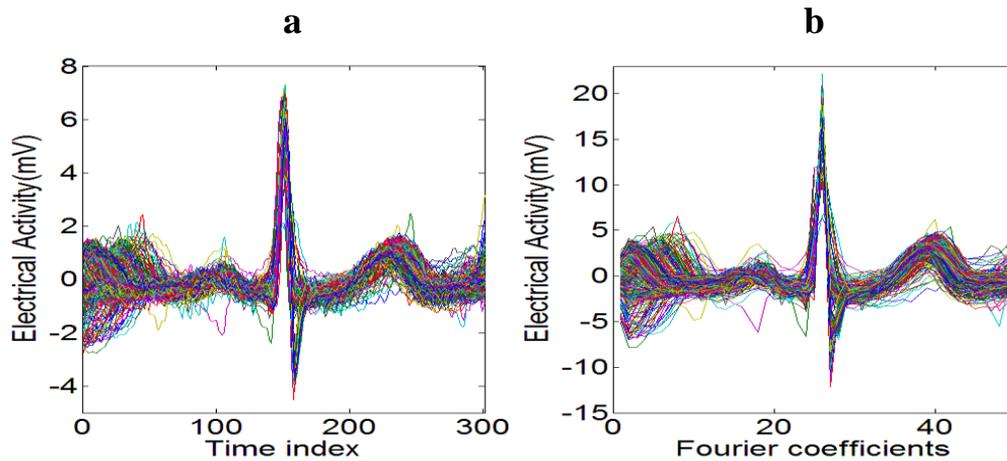


Figure 5. 20 : Extract of (A) NORMAL, (B) PVC, (C) APC and (D) FUSION beats.

Both Discrete Cosine Transforms (DCT) and Discrete Sine Transforms (DST) were used for feature extraction and dimensionality reduction of the input vector to the multi-class SVM classifier as this operates only with real numbers. However, Discrete Fourier Transforms (DFT) were used for feature extraction and dimensionality reduction of the input vector to the CSVM classifier which can handle complex valued datasets. Between 30 and 50 DFT coefficients were selected in the frequency domain for reconstructing individual ECG beats in the feature selection phase. A similar number of DST and DCT coefficients were also used. Fig 5.22 shows a comparison between original normal beat feature and reconstructed normal beats from only 50 Fourier coefficients showing minimal distortion.



**Figure 5.22:** (a) Typical ECG datasets extracted with 301 sample window (b) time domain filtered and reconstructed signal using 50 Fourier coefficients.

For the classifier implementation, after feature selection is performed, the datasets are divided into two groups with 500 beats for training and 500 beats for testing purposes. Each of these datasets contain 200 NORMAL, 100 PVC, 100 APC and 100 FUSION beats. As can be seen the number of beats in each class were increased from 205 to 500 compared to the previous study in order to measure the performance of the algorithm with that using a larger dataset. Sequential minimal optimization (SMO) [188] was used to compute hyper-plane parameters and threshold values for both MSVM and CSVM during the training phase. The optimal hyper-plane parameters are identified on the basis of the training data set, and class label. Finally, on the basis of the training results, test data are imported to the CSVM and MSVM classifiers to perform unknown beat classification. The LIBSVM software is used to train and validate the MSVM model and to perform classification. The performance of the classifiers is evaluated by computing of the ST, SP, and PP parameter and the total

classification accuracy. Classification results of correct and misclassified beats for each class were displayed by a confusion matrix in Tables 5.8 and 5.9. When an ECG beat is misclassified as another one, this can be identified from using the output of the confusion matrix. Table 5.8 presented the confusion matrix of the classification results using CSVM, while the classification results using MSVM are shown in Table 5.9. It can be seen that all PVC beats are correctly classified when using both algorithms (MSVM and CSVM). It was found that with MSVM, 20 NORMAL beats 33 PVC beats and 30 FUSION beats were misclassified as an APC beat, while with CSVM only one NORMAL beat was misclassified as APC beats. Two NORMAL beats and two PVC beats were classified as FUSION beats using CSVM, whereas nine NORMAL beats and three PVC beats were misclassified as FUSION beats.

Table 5.10 summaries the results of the classification process using the three common measures: SE, SP, and PP. The proposed algorithm achieved classification accuracies up to 97%, whereas multi-class SVM achieved up to 83% accuracy. The sensitivities were found to be 99%, 100%, 100%, 97% and 100.0% for NORMAL, PVC, APC and FUSION cases, respectively using CSVM. The sensitivities were 95.59%, 91.32%, 90.50%, 94.51%, and 93.77% for NORMAL, PVC, APC and FUSION beats cases respectively using MSVM.

Annotation	Output result			
	NORMAL	PVC	APC	FUSION
NORMAL	190	0	1	9
PVC	2	95	0	3
APC	0	0	100	0
FUSION	1	0	0	99

**Table 5.8:** Confusion matrix classification results using DFT extracted coefficients of the ECG signal taken from a single limb lead and using CSVM.

Annotation	Output result			
	NORMAL	PVC	APC	FUSION
NORMAL	178	0	20	2
PVC	0	65	33	2
APC	0	0	100	0
FUSION	0	0	30	70

**Table 5.9:** Confusion matrix classification results using DFT extracted coefficients of the ECG signal taken from a single limb lead and using MSVM.

Annotation	Classification performance result					
	CSVM			MSVM		
	PP (%)	ST (%)	SP (%)	PP (%)	ST (%)	SP (%)
NORMAL	98	95	99	100	89	100
PVC	100	95	100	100	65	100
APC	99	100	100	55	100	79
FUSION	89	99	97	95	70	99

**Table 5.10:** Collective performance analysis and classification result using MSVM and CSVM in the study in single limb lead study presented in section 5.5.2.2.

Using CSVM showed an improvement in the sensitivity for all classes and significantly enhanced the overall classification accuracy. However, when comparing these results with those in the precordial lead study when the number of beats for training and testing was increased, a slightly decreased classification accuracy was observed. In addition, it is worth noting that the ECG signals used in these studies were obtained using one limb lead (II) and two patients records, whereas in the precordial lead study the ECG beats were obtained from four precordial leads and 12 patient records. Extracting beats from more than one lead while using several patient records could improve classification accuracy results as was shown in the previous study.

This work was presented at the 9<sup>th</sup> Saudi Students Conference in the UK 2016 as a poster and was published in [29].

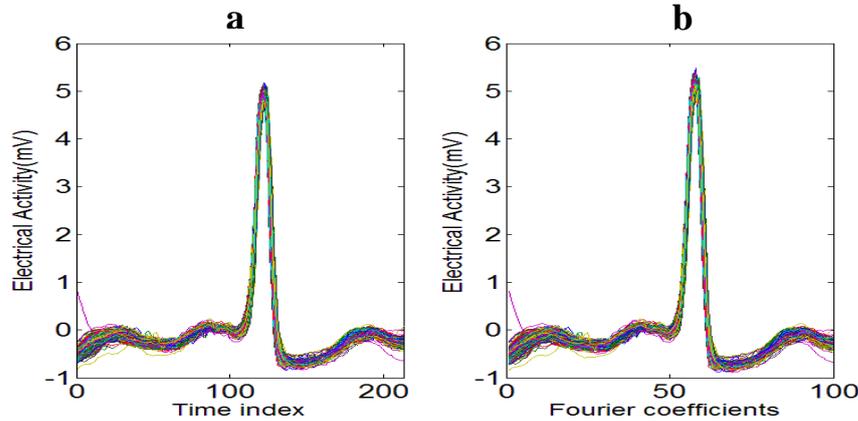
### **5.5.2.3 Results from a MSVM and CSVM study using two leads (V1, V5) with data from the European ST-T database identifying NORMAL, PVC, SVPB and FUSION beats**

In this section, an automatic classification of four beat types NORMAL, PVC, Supraventricular premature or ectopic beats SVPB and FUSION of ventricular and normal beats is discussed, as implemented using the MSVM and CSVM algorithms.

The ECG signals used in these experiments were obtained from the European ST-T Database using recordings from two leads (V1, V5). A number of beats from different patients were selected for training and evaluating classifier performance. In the pre-processing stage, baseline removal and R peak detection were performed. R peak detection is performed with the WFDB Software Package. R (peak) location was used as a reference to detect peaks in other waves such as P and T and to extract the ECG beat. ECG beats are extracted after windowing the signal using 106 samples before the R-peak and 106 samples after the R-peak as mentioned in a previous study [209].

Discrete Cosine and Sine transforms or the Discrete Fourier Transform (DFT) were used for feature extraction and dimensionality reduction of the input vector at the input of the classifier. Studies after selecting either 100 or 50 Fourier coefficients for reconstructing individual ECG beats in the feature selection phase were performed. Using 100 Fourier coefficients indicated better result than using 50 Fourier coefficients due to smaller signal distortion in the reconstruction process. These were used for creating the input vector to the classifier and reconstructing individual ECG beats as shown in Figure 5.23. DCT and DST

coefficients were used to extract the features presented at the input vector of the MSVM classifier whereas DFT coefficients are used for creating an input vector to the CSVM classifier. Fig 5.23 illustrates a comparison between the normal ECG beats with 213 samples around the R-peak at a sampling frequency of 250 Hz and reconstructed normal ECG beats using only 100 Fourier coefficients.



**Figure 5.23: (a) Typical ECG datasets extracted with 213 sample window (b) time domain filtered and reconstructed signal using 100 Fourier coefficients.**

The MSVM and CSVM classifiers are used to distinguish between the four ECG arrhythmias types using the DCT, DST and DFT coefficients as feature vectors at the input vector of the classifier. For the classifier implementation, after feature selection is performed, the datasets (622 beats) are divided into two groups with 311 beats for training and testing purposes respectively. Each of these group contain 150 Normal beats (NORMAL), 60 premature ventricular contraction (PVC) beats, 60 Supraventricular premature or ectopic (SVPB) beats and 41 Fusion of ventricular and normal (FUSION) beats. MATLAB software routines were used to train and validate both the CSVM and the Multi-class Support Vector Machine (MSVM) classifier, with LIBSVM used to train and validate MSVM. Firstly, the kernel and complex kernel functions (Gaussian RBK) were used and 5-fold cross validation was adopted for adjusting the kernel and SVM parameter values. Then, SMO is used to train the CSVM and compute the corresponding complex hyper-plane parameters as in the previous studies. LIBSVM software routines are used to train and validate the MSVM model using SMO as well as the kernel parameter values that were obtained using 5-fold cross validation. In both CSVM and MSVM, the optimal hyper-plane parameters are identified on the basis of the training data set and class label. The output of the SMO algorithm was used for training. Finally, on the basis of the training results, test data are imported to the CSVM and MSVM classifiers to perform unknown beat classification. The optimal valued for  $a_n$ ,  $b_n$  and

threshold were adopted and applied to the test datasets in order to measure the classification performance for both CSVM and MSVM classification.

Classification accuracies of up to 94% were obtained using CSVM, whereas multi-class SVM achieved up to 86% accuracy. Confusion matrices were used as part of the evaluation to summarise the classifier performance. These are shown in Tables 5.11 and 5.12. As can be seen in Table 5.11 three of the 60 PVC beats are misclassified as SVPB whereas four are estimated as misclassified FUSION beats. Five of the 41 FUSION beats are misclassified as SVPB, while six are misclassified as either three NORMAL beats or three PVC beats. 150 NORMAL beats and 60 SVPB beats are correctly classified. According to the overall confusion matrix given in Table 5.12, one of the 150 NORMAL beats is identified as PVC while six the 60 SVPB beats are estimated as PVC beats. One of the 60 PVC beats is misclassified as a FUSION beat, whereas one of the 41 FUSION beats is identified as an SVPB beat. In addition, from the confusion matrix, the sensitivity and specificity of the implemented algorithm for individual classes is obtained. Table 5.13 illustrates the performance of the classification process using the three metrics (SE, SP, and PP). This table included the final decision results obtained using the two method: CSVM and MSVM. The sensitivities were found to be 100%, 88%, 100%, and 73% for NORMAL, PVC, SVPB and FUSION beats respectively using CSVM. In contrast, using MSVM differed sensitivities were obtained: 99%, 98%, 90%, and 17% for NORMAL, PVC, SVPB and FUSION beats respectively. It may thus be concluded that using CSVM a significantly improved classification accuracy is obtained. This is better than using MSVM, which produced sensitivities between 86% to 94%.

Annotation	Output result			
	NORMAL	PVC	SVPB	FUSION
NORMAL	150	0	0	0
PVC	0	53	3	4
SVPB	0	0	60	0
FUSION	3	3	5	30

**Table 5.11: Confusion matrix using MSVM and European ST-T database records.**

Annotation	Output result			
	NORMAL	PVC	SVPB	FUSION
NORMAL	149	1	0	0
PVC	0	59	0	1
SVPB	0	6	54	0
FUSION	0	33	1	7

**Table 5.12: Confusion matrix using CSVM and European ST-T database records.**

Annotation	Classification performance result					
	CSVM			MSVM		
	PP (%)	ST (%)	SP (%)	PP (%)	ST (%)	SP (%)
<b>NORMAL</b>	98	100	98	100	99	100
<b>PVC</b>	95	88	99	60	98	84
<b>SVPB</b>	88	100	97	98	90	100
<b>FUSION</b>	88	73	99	88	17	100

**Table 5.13: Collective PP, ST and SP results using MSVM and CSVM classifiers for a 2-lead study using the European ST-T database records.**

Results illustrate that the proposed beat classifier is very reliable, and that it may be adopted for automatic detection of arrhythmia conditions and classification. Moreover, the present research confirmed that the use of selected number of Fourier coefficients to approximate the ECG beat signal and compress the input features to the classifier could lead to high classification accuracies and improve the generalization ability of the CSVM classifier. To our knowledge, this is the first time that multiclass algorithms such as MSVM and CSVM were employed to classify ECG beats that were extracted from the European ST-T Database.

This work was presented at 2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB) conference, Temple University in the USA as a poster and was published in [28].

#### **5.5.2.4 Discussion of results from the second study using the MSVM and CSVM classifiers**

Classification of arrhythmias is a complex problem because of the strict requirement for avoiding false-positive or false-negative results. Pre-processing and feature extraction are crucial steps for developing automated diagnostic expert systems for different types of heart disease. The high classification accuracy of the multiclass CSVM gives additional insights into the features that may be selected for identifying the ECG signals.

In addition, it is worth noting that as discussed in chapter 3 there are several alternative classifiers that can be used in an ECG classification context, such as a linear discriminant analysis (LDA) classifiers and artificial neural networks (ANNs). There have been suggestions to use novel hybrid neural network [15], probabilistic neural network [10], back propagation neural networks, self-organizing maps (SOM), learning vector quantization (LVQ) schemes [108], support vector machines (SVM) [210] and fuzzy or neuro-fuzzy algorithms [109]. In order to place our results in context to these studies, a performance comparison of the proposed CSVM classifier to other classifiers based on traditional SVM

and Least Square Support Vector Machines (LSSVMs) is shown in Table 5.14. The table illustrates the different approaches in terms of the chosen feature selection method, number of features used as input vector to the classifier, number of arrhythmia types studied and overall accuracy. Moreover, it focuses on multi-class SVM studies. For brevity, it excludes binary SVM classification studies.

Martis *et al.*, (2012) indicated that through the PCA of time domain ECG signals, feed-forward neural network (NN) and Least Square- Support Vector Machine (LS-SVM) can be used to classify five types of ECG beats (NORMAL, RBBB, LBBB, APC and PVC) with an accuracy of 98.11% [137]. In another study, fast least square support vector machines (LS-SVMs) classification and DCT were used to classify six types of ECG beats (NORMAL beat, LBBB beat, congestive heart failure beat, PVC beat, non-conducted P-wave, ventricular escape beat) with a 95.2% rate of accuracy [24]. Rabee and Barhumi (2012) used SVM with wavelet multi-resolution (through the discrete wavelet transform) pre-processing for classifying 14 different types of heart beats. Their work was based on a comprehensive study that took into consideration, 17260 ECG beats, selected from the MIT/BIH database. The average classification accuracy achieved was 99.2% [90]. Moreover, extracted principal components of the bi-spectrum using an LS-SVM with an RBF kernel were used to classify the same five types of beats with a 93.48% rate of accuracy [20]. In another study, SVM in conjunction with multi-class directed acyclic graphs was used to classify four types of ECG signals (NORMAL beat, Atrial fibrillation, Ventricular tachyarrhythmia and Congestive heart failure). The features of input vectors were selected using empirical mode decomposition and singular value decomposition. The Directed Acyclic Graph Support Vector Machine (DAGSVM) classifier yielded an average accuracy rate of 98.96%. In addition, the DAGSVM algorithm and two other classifier methods (K-Nearest Neighbor (KNN) and Artificial Neural Network (ANN)) were also compared. DAGSVM showed an accuracy of 98.96%, while the average accuracy of the classification result using KNN and ANN algorithms was 95.83% and 96.66% respectively [224].

In studies by Martis *et al.*, (2013) five types of beat classes of arrhythmia were analyzed, namely non-ectopic beats, supra-ventricular ectopic beats, ventricular ectopic beats, fusion beats and unclassifiable and paced beats. It was found that the DWT can provide acceptable time and good frequency resolutions for ECG datasets. In their study, The SVM, neural network (NN) and the PNN were used for classifying these five types of ECG beats for automated diagnosis [217].

Feature reduction method	Number of feature	Classification method	Number of arrhythmia type	Overall Accuracy (%)	Reference
PCA	30 PC	SVM	4	99.17	[136]
PCA	20 PC	SVM	4	99.08	[136]
PCA	10 PC	SVM	4	98.69	[136]
DFT	15 Fourier coefficients	LS-SVM	6	92.2	[24]
DCT	15 DCT coefficients	LSSVM	6	95.2	[24]
DWT	15 Wavelet coefficients (WFs)	LS-SVM	6	94.2	[24]
Adaptive autoregressive (AAR)	15AAR coefficients	LS-SVM	6	91.7	[24]
DWT	16 WFs	SVM	4	94	[211]
DCT	18 DCT coefficients	SVM	4	96.5	[211]
average of every eight samples of 256 (Amplitude value)	32	SVM	4	94.2	[211]
DWT	20 WFs	Multi-class SVM	4	98.61	[101]
ICA	20IC	SVM	6	98.7	[212]
GDA	5	SVM	6	99.16	[104]
PCA	15	SVM	6	97.65	[104]
LDA	15	SVM	6	98.06	[104]
Genetic Algorithm	22	SVM	4	93.46	[213]
PCA	22	SVM	4	80.00	[213]
ICA	17 ICs	SVM with	8	98.7	[13]
DWT	25 WFs	Multi-class SVM	7	98	[135]
Non-parametric power spectral density (PSD)	132	SVMGA	5	96.00	[5]
Cross-correlation & FFT	32	LS-SVM	3	95.82	[214]
DWT & PCA	20 PC	SVM	4	99.63	[215]
PCA	12 PC	LS-SVMwith RBF kernel	5	98.11	[137]
Linear Prediction (LPC) & PCA	12 PC	LS-SVMwith RBF kernel	5	94.88	[137]
DWT& PCA	12 PC	LS-SVM with RBF kernel	5	96.88	[137]
DWT	251	SVM with GRBF	14	99.2	[58]
DWT and Statistical with electrophysiological features	68	Multi-class SVM	12	98.92	[216]
Higher Order Statistics (HOS) ( Bispectrum) & PCA	12 PC	LS-SVM	5	93.48	[20]
Continuous WT	129	SVM	6	99.82	[14]
WT	17	SVM	6	98.74	[17]
WT+LDA	4	SVM	6	99.52	[17]
WT+PCA	4	SVM	6	98.86	[17]
DWT& LDA	12 LDA features	SVM with RBF kernel	5	97.04	[217]
DWT& PCA	12 PC	SVM with RBF kernel	5	96.92	[217]
DWT& ICA	12 ICA	SVM with RBF kernel	5	98.36	[217]
DCT and PCA	12 PC	LS-SVM with RBF kernel	5	96.61	[218]
ECG Chaos Extractor platform	11 (four linear and seven nonlinear HRV features)	SVM	4	98.4	[219]
statistical correlation analysis	11	SVM	11	97.87	[220]
autoregressive (AR)	300	SVM	5	99.93	[221]
a geometric technique ( <i>Temporal and amplitude distance</i> )	21	SVM	15	98.80	[222]
DWT	17	SVM	8	98.42	[223]
SVD	9	DAGSVM	4	97.71	[224]
ICA	200	SVM	3	99.9	[225]
DFT	100	CSVM	4	94	Current work [28]
DFT	50	CSVM	4	97	Current work [29]
DFT	50	CSVM	4	100	Current work

**Table 5.14: Comparison of overall accuracy of different ECG classifier reported in literature addressing the multi-class classification problem.**

Moreover, the performance of this study was compared with other well-established methods of ECG classification such as Combined Neural Networks (CNNs), Radial basis function neural network (RBFNN) and neural network.

Guler and Ubeyli (2005) showed that four types of ECG beats (NORMAL beat, congestive heart failure beat, ventricular tachyarrhythmia beat, and atrial fibrillation beat) were successfully classified with a 96.94% rate of accuracy by using Combined Neural Networks (CNNs) when trained with the extracted features from discrete wavelet transforms of NORMAL beat, congestive heart failure beat, ventricular tachyarrhythmia beat, and atrial fibrillation beat signals [129]. The same group also used WTs for feature extraction and an adaptive neuro-fuzzy inference system (ANFIS) trained with backpropagation to classify two NORMAL beats and partial epilepsy beats, with an accuracy rate of 98.13% [226].

In (2010) Korürek and Doğan illustrated classification using a Radial basis function neural network (RBFNN) for classifying six types of ECG beats, namely NORMAL beats, PVC beats, FUSION beats, APC beats, RBBB beats and PFUS beats) [113]. In addition, the PNN and a backpropagation neural network were used with IPCA to differentiate between eight different ECG beat types with an accuracy above 98%. The IPCA was used to extract features from ECG signals and create input vectors for the classifier [9].

Departing from the standard neural network methodologies, in 2006, Polat *et al.*, proposed that fuzzy weighted pre-processing and artificial immune recognition system (AIRS) can be also used for pre-processing ECG datasets and diagnose cardiac arrhythmias. It was found that 78.79%, 75.00% and 80.77% classification accuracies were obtained using 50-50%, 70-30%, and 80-20% of training dataset to validation dataset ratios respectively [227]. In addition, Ceylan *et al.*, in 2009, built a module that contained a combination of a type-2 fuzzy c-means (T2FCM) clustering algorithm along with a neural network, for the classification of ten types of ECG arrhythmias obtained from the MIT-BIH database (normal beat, sinus bradycardia, ventricular tachycardia, sinus arrhythmia, APC, paced beat, RBBB, LBBB, atrial fibrillation and atrial flutter). This last study was further refined with the addition of a wavelet transform feature extraction, to obtain input vectors for the classifier [102].

Luz *et al.*, (2014), proposed a SVM classifier using four feature selection methods, where WT, and IPCA was also used to classify ECG signals of low frequencies (sampled at

30 and 60 Hz respectively) with an accuracy of 95% [228]. In addition, five types of ECG beats were classified, using an alternative pattern recognition technique known as the Optimum-Path Forest (OPF) classifier. The performance resulting from the use of OPF was compared with other classification methods such as SVM, Bayesian and MLP models [229].

The multiclass SVM results from the current CSVM study compares favourably to most of the above studies. Using CSVM showed slightly better performance than the multiclass SVM classifiers that were used in all previous studies. The results are very encouraging when considering four classes of ECG beats were simultaneously classified. In addition, using the ECG beat from two leads showed a significant improvement in the classification accuracy as compared to the above classifiers, achieving an impressive average accuracy of 100%. This is an important finding justifying the use of both leads in clinical practice.

Finally, it is also interesting to note that the SVM Kernel-Adatron (K-A) learning algorithm and the backpropagation (BP) learning algorithm with a multi-layered perceptron (MLP) have also been used to classify six different types of arrhythmias and normal ECG beats, without resorting to any feature extraction pre-processing methods [230]. Despite that result, in the current study it was maintained that feature extraction through pre-processing is absolutely necessary for best classification performance.

In future investigations, we intend to use twelve different leads, where several types of beats from pathogenic and normal datasets are simultaneously selected. Using ECG signals from multiple leads provides a more complete patient picture and enables simultaneous identification of various heart conditions. The proposed approach can be extended using an adaptive wavelet bank for pre-processing (see appendix) where a different wavelet function is derived at each decomposition level to improve the parsimony of the input vector [206].

Furthermore, multidimensional SVMs using Clifford Algebras [231] should also be considered to account for multi-lead signals as well as correlate inputs from alternative sensing modalities. Such approach will make the proposed classification methodology more generic and relevant to future advances in sensing.

### **5.5.3 Study three: SIMCA classifier structure and results**

SIMCA is a method that is widely adopted to classify multiple objects. This model helps in finding similarities between test datasets and classes rather than matching identical behaviour.

SIMCA modelling is based on the PCA. The following study looks at the use of a new ECG arrhythmia classification scheme based on PCA for feature extraction using a SIMCA classifier to differentiate between six types of arrhythmia conditions. The aim of this work is to improve the classification accuracy when using a large database composed of different classes of pathological conditions and different beat duration. The previously described data pre-processing and beat extraction methodology is adopted, feature extraction and selection and a classification step as shown in Fig 5.24 which depicts the three most important steps in the classification process is adopted.

### **5.5.3.1 ECG acquisition and pre-processing for the SIMCA study**

Six types of beats including: NORMAL, PVC, APC, RBBB, LBBB and FUSION beats are selected from the MIT-BIH arrhythmia database. A number of beats of these six type were selected from limb lead (II), precordial leads (V1), these related to different patients. All these records were presented simultaneously to the classifier for training and for evaluating classifier performance. The research conducted focuses on ECG beat segments that include the P, QRS complex and T wave as in the previous studies. Again first the R peak location needs to be identified using the annotation files so, the associated RR interval can be found from the locations of the R points documented in the annotation files. Each complete beat is time domain a segment that has 127 data points before and 128 data points after the R evens. In each time domains sequence there are 256 points corresponding to a sampling frequency of 360 Hz. A rectangular window (256 data points) is formed by centring the R peak in the QRS complex for each ECG beat segment as determined in the previous studies [24],[129]. For the current study, nine recordings, with six ECG beat types from two leads where used. Table 5.14 illustrates the details of the two dataset groups used in terms of record number, lead name and total number of beat samples for each beat type studied. ECG extracted beats need to be pre-processed and baseline corrected before extracting the features and creating the input vector to the classifier. Fig 5.25 illustrated the extraction of six types of beat with 256 samples around the R peak with normalized signals with zero mean and unity standard deviation. In order to improve classification performances the dimensions of the beat segments need to be reduced using feature reduction and selection techniques such as PCA.

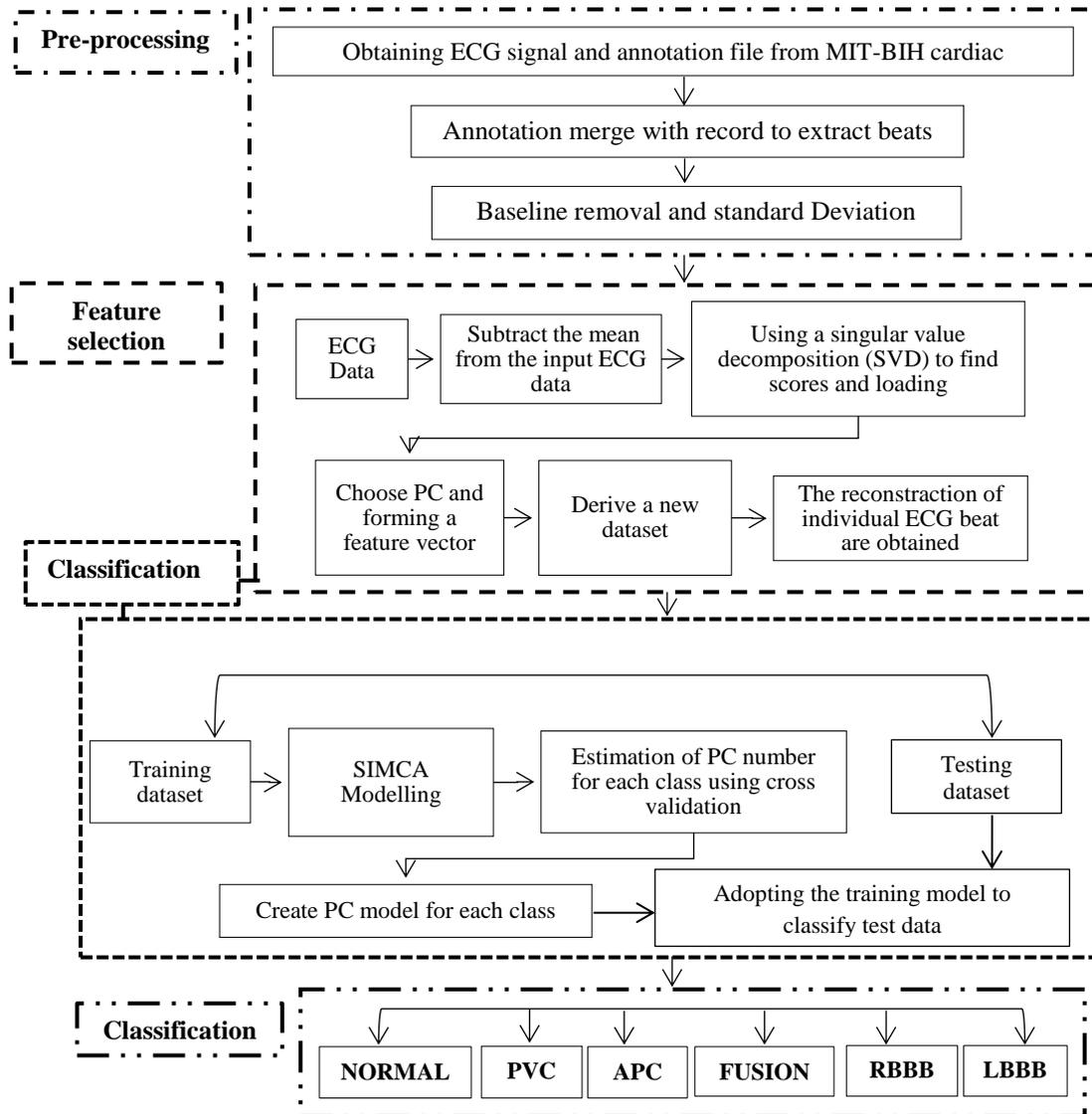


Figure 5.24: Block diagram of proposed method for the SIMCA cardiac arrhythmia classification study.

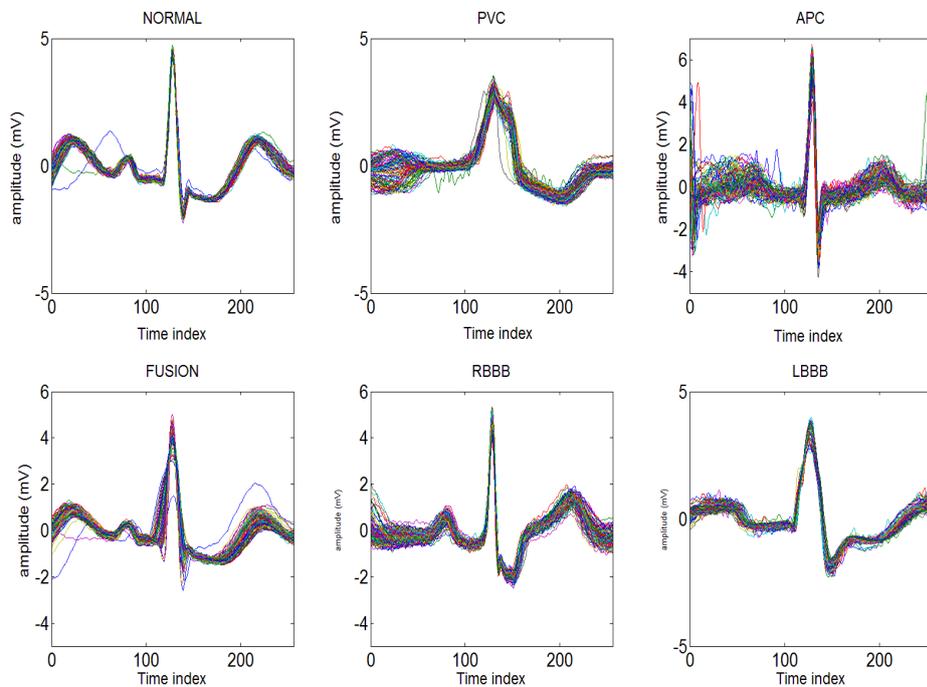
### 5.5.3.2 Feature extraction and calculation of the feature vectors

As it was mentioned the ECG features can be extracted either in the time [92] or the frequency domain [155]. The features of the P, QRS complex, and T waves are associated with the location, duration, amplitudes, and shapes of the waves. There are several dimensionality reduction and feature extraction algorithms that may be applied successfully to ECG signals such as PCA [10], LDA [217] and wavelet transforms (WT) [24]. PCA is a well-established technique for feature extraction from ECG beat and dramatically reduce dimensionality of input vector to classifier without assuming a particular approximation

function in the traditional sense by relying on the statistical similarities between the time domain sequences.

Dataset	Annotation	Record #	Lead Group	
			Limb lead	precordial lead
			II	V1
Train	NORMAL	209, 208, 213, 223	200	200
	PVC	208, 213, 223, 214	200	200
	APC	201, 202, 223, 214	200	200
	FUSION	208, 213	200	200
	RBBB	118, 212	200	200
	LBBB	111, 109, 214	200	200
Test	NORMAL	209, 208, 213, 223	200	200
	PVC	208, 213, 223, 214	200	200
	APC	201, 202, 223, 214	200	200
	FUSION	208, 213	200	200
	RBBB	118, 212	200	200
	LBBB	111, 109, 214	200	200
Total		11	2400	2400

**Table 5.14 Illustration of two ECG beat groups from the MIT-BIH cardiac arrhythmia database used in both training and test datasets.**



**Figure 5.25: Extract of six types of ECG beats from the MIT-BIH cardiac arrhythmia database.**

In (2007) Ceylan and Özbay used four approaches simultaneously: a technique for fuzzy c-means clustering (FCM), Principal component analysis, FCM with PCA and WT to perform feature extraction and data reduction; their study focused on developing a classifier for ten different arrhythmias conditions [175]. In another study [175] different structures including PCA such as PCA-NN and fuzzy c-means clustering (FCM)-PCA-NN were used for the classification of ECG arrhythmias. PCA was applied to reduce sample number in each beat from 200 to 20. The study concluded that the FCM-PCA-NN structure showed better result than the PCA-NN structure. In 2007 Polat and Gunes proposed the use of PCA to decrease the ECG Arrhythmias beat features from 279 features per beat to 15. It was found that 100% test classification accuracies were obtained using these features [174]. In the present study, PCA was used to reduce the dimensions of the normalised ECG beat samples. PCA is a well-established technique for feature extraction and data compression without much loss of important information. Between 10 and 30 principal components were selected for each beat as input features for classification in order to determine the optimal number of PCs. The approach of justifying these numbers and the method used for finding optimal number of features is based on extensive cross validation as discussed in the next section.

### **5.5.3.3 Classification performance with SIMCA**

For each experiment 2400 beats of six ECG beat types taken from both precordial lead (V1) and limb lead (II) from the MIT-BIH arrhythmia database and were selected. Each dataset includes 400 N, 400 PVC, 400 APC, 400 RBBB, 400 LBBB, and 400 FUSION segments. The ECG beat samples were randomly divided into two subsets. Half of the beat samples were selected as the training subset and the other half as the testing subset. Mean centered normalised ECG beat samples were used to reduce dataset and better numerical on denoising of the SVD associated with the PCA. In order to estimate the best number of the principle components (PCs) of the PCA, various combinations of PC number were used to reconstruct the ECG beat samples and used as input to the classifier. The feature reduction method for the input vectors was repeated four times in order to create feature vectors of different length for ECG beats with an optimal number of principle components (PCs). The best number of components in each class could be determined through the cross-validation process as well. PCA is performed first on each group of classes and then SIMCA is used on the PCA classes. PCA provides a matrix of scores and loadings for each class. In addition, each class can be summarized in a different dimension. Fig 5.26 and 5.28 illustrated the predicted residual error

sum of squares (PRESS) statistic curve for the six classes using cross validation. The optimal number of the principal components (PCs) was used to reconstruct individual ECG beats in the dataset. Once the optimal dimensions for each beat were obtained, the classification performance of the SIMCA classifier was evaluated for solving the problem of identifying the six different types of ECG arrhythmias. The performance of SIMCA is optimized by creating individual models for each class and finding similarities between test objects and class models. The first step in SIMCA modelling is therefore to build a model for each class using an optimal number of principal components (PCs) and creating a training dataset. Individual models for each class were created using the optimal number of the principal components (PCs). These models were combined with the test dataset to assess the classification performance attained.

To evaluate the performance of the SIMCA classifier ECG beats from Limb lead (II) and chest lead (V1) were grouped in two respectively and presented at the input of the classifier individually. Limb lead (II) dataset is firstly used, and then a second dataset with ECG beats obtained from precordial lead (V1) was used. The classification performance was evaluated after presenting 2400 beats associated with the six ECG beat types. These were related to different patient records as shown in Table 5.14. Patients records were randomly divided into four equally sized groups consisting of 2400 beats each for training and testing. To evaluate the performance of the proposed arrhythmia classification algorithm, three common measures of SE, SB, ACC and predictively PP are used. Confusion matrices are shown in the following section.

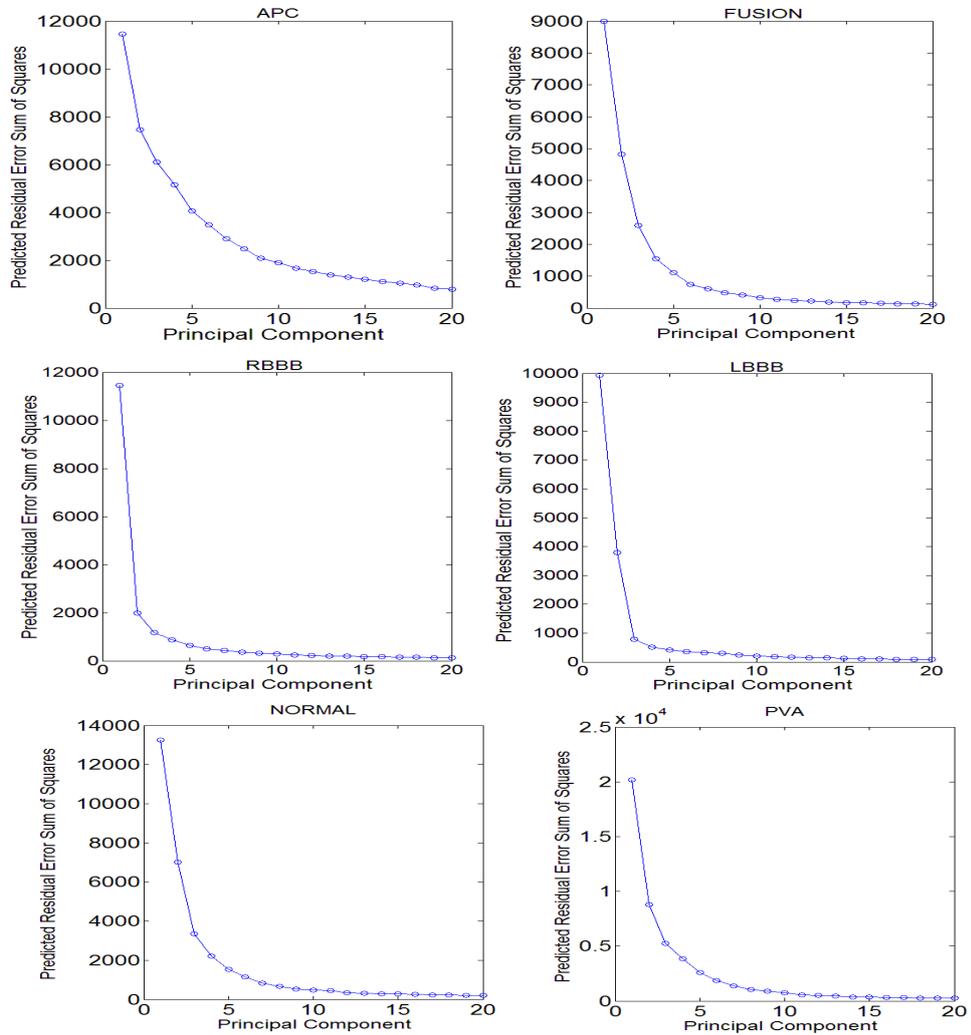
#### **5.5.3.4 SIMCA classification using the limb lead datasets from the MIT-BIH Arrhythmia Database**

Fig 5.26 illustrated the predicted residual error sum of squares (PRESS) curve for the six classes using cross validation. It can be seen that the average predicted residual error sum of squares reached a plateau at around 10 PCs in the majority of classes. There was no noticeable improvement when using more PCs. In addition, Table 5.15 summaries the model accuracy for different number of PCs used as input to SIMCA. As can be seen, the best accuracy of 95.83% was obtained using 10 PCs. Reducing the number of features to ten will increase the number of misclassifications in some sample classes. Therefore, the dimension of each ECG beat sample was reduced from 256 to 10 using the PCA method and an optimal number of principal components (PCs). This is a very impressive result, showing the

advantages of the proposed methodology. It is worth stating that using only 10 PCs is also beneficial from the perspective of not overwhelming the classifier so that its generalization capability can be maximized.

PCs number	Model accuracy
10	98.61
15	98.33
20	98.22
30	97.28

**Table 5.15: Classification accuracy rate for different number of PCs coefficients presented to classifier using ECG beats obtained from limb lead (II).**



**Figure 5.26: predicted residual error sum of squares (PRESS) curve for the six beat type classes using ECG beats obtained from the limb lead.**

Moreover, the classification results for each ECG beat type are displayed using confusion matrices, which illustrate the success of the classification process in terms of classified or misclassified results. The confusion matrix of the ECG arrhythmia classification using the limb lead datasets and SIMCA are shown in Tables 5.16. It can be seen that four of the NORMAL, one PCV and three of the LBBB beats were determined as APC, while 11 APC beats were estimated as Normal. All of the RBBB beats were correctly classified. A total of 197 of the 200 RBBB beats were correctly classified, and three were identified as APC beats. Seven of the PVC beats and only one of the NORMAL beats were determined to be FUSION beats, whereas 14 of 200 FUSION beats were classified as NORMAL and eight as PVC.

Annotation	Output result					
	NORMAL	PVC	APC	FUSION	RBBB	LBBB
NORMAL	195	0	4	1	0	0
PVC	1	191	1	7	0	0
APC	11	0	189	0	0	0
FUSION	14	8	0	178	0	0
RBBB	0	0	0	0	200	0
LBBB	0	0	3	0	0	197

**Table 5.16: Confusion matrix of the ECG signal taken from limb lead and SIMCA classification.**

The overall performance of the method for test dataset in terms of SE, SB, PP and ACC with SIMCA is illustrated in Tables 5.17. The best and poorest classification sensitivity of each ECG beat type were 100% for RBBB and 89% for FUSION beats respectively whereas the best and poorest SP were 100 % for RBBB and LBBB beats and 97.40% for NORMAL beats respectively. The sensitivities were 97.50%, 95.50%, 94.50%, 89%, 100 % and 98.50% for the NORMAL, PVC, APC, RBBB and LBBB heartbeat classes respectively. The SIMCA classification algorithm could simultaneously classify six beat types obtained from limb lead, with an average accuracy of 98.61%.

Annotation	Classification performance result			
	PP (%)	ST (%)	SP (%)	ACC (%)
NORMAL	88.24	97.50	97.40	97.42
PVC	95.98	95.50	99.20	98.58
APC	95.94	94.50	99.20	98.42
FUSION	95.70	89	99.20	97.50
RBBB	100	100	100	100
LBBB	100	98.50	100	99.75

**Table 5.17: Collective result performance analysis and classification result using beats obtain from limb lead (II).**

Tables 5.16, 5.18, 5.19 and 5.20 illustrate the collective result of classification for each class individually versus using a different number of PCs as input to the SIMCA classifier. The most accurate performance for each class was obtained for NORMAL beats with 97.50% sensitivity, PVC with 95.50 %, RBBB with 100%, LBBB with 98.50%, and 10 PCs; while APC with 100% and 20 PCs, and FUSION with 91.00% sensitivity and 15 PCs. With the use of 10 PCs, 22 of the 200 FUSION beats are estimated incorrectly, while 18 of the 200 FUSION were misclassified using 15 PCs. All of the 200 APC beats were correctly classified using 20 PC, whereas 11 APC were determined as NORMAL using 10 PCs. Using 15 PCs showed an improvement in classification sensitivity of APC. Moreover, 20 PCs illustrated an improvement in classification sensitivity and accuracy of FUSION beats. Using 15 PCs and 20 PCs showed an improvement in the classification performance of the APC and FUSION beats respectively. Fig 5.27 illustrates the collective result of classification accuracy for each class individually versus using a different number of PCs as input to the SIMCA classifier.

Annotation	Output result					
	NORMAL	PVC	APC	FUSION	RBBB	LBBB
<b>NORMAL</b>	181	0	19	0	0	0
<b>PVC</b>	1	186	5	8	0	0
<b>APC</b>	0	0	200	0	0	0
<b>FUSION</b>	13	7	1	179	0	0
<b>RBBB</b>	0	0	0	0	200	0
<b>LBBB</b>	0	0	10	0	0	190

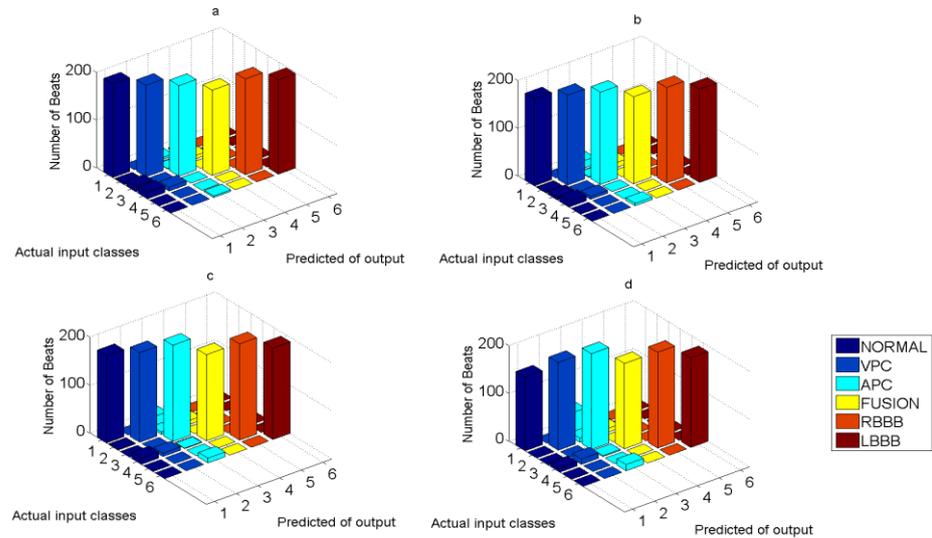
**Table 5.18: Confusion matrix of the ECG signal taken from the limb lead using SIMCA classifier and 15 PCs.**

Annotation	Output result					
	NORMAL	PVC	APC	FUSION	RBBB	LBBB
<b>NORMAL</b>	154	0	44	2	0	0
<b>PVC</b>	0	183	9	8	0	0
<b>APC</b>	0	0	200	0	0	0
<b>FUSION</b>	9	11	2	178	0	0
<b>RBBB</b>	0	0	0	0	200	0
<b>LBBB</b>	0	0	13	0	0	187

**Table 5.19: Confusion matrix of the ECG signal taken from the limb lead using SIMCA classifier and 20 PCs.**

Annotation	Output result					
	NORMAL	PVC	APC	FUSION	RBBB	LBBB
<b>NORMAL</b>	182	0	15	3	0	0
<b>PVC</b>	1	188	1	10	0	0
<b>APC</b>	7	0	193	0	0	5
<b>FUSION</b>	12	6	0	182	0	0
<b>RBBB</b>	0	0	0	0	200	0
<b>LBBB</b>	0	0	5	0	0	195

**Table 5.20: Confusion matrix of the ECG signal taken from the limb lead using SIMCA classifier and 30 PCs.**



**Figure 5.27: the predicted of input individual ECG beat to the SIMCA classifier ECG beats taking from limb lead (V1) using (a) 10 PCs , (b) 15 PCs, (c) 20 PCs and (d) 30 PCs.**

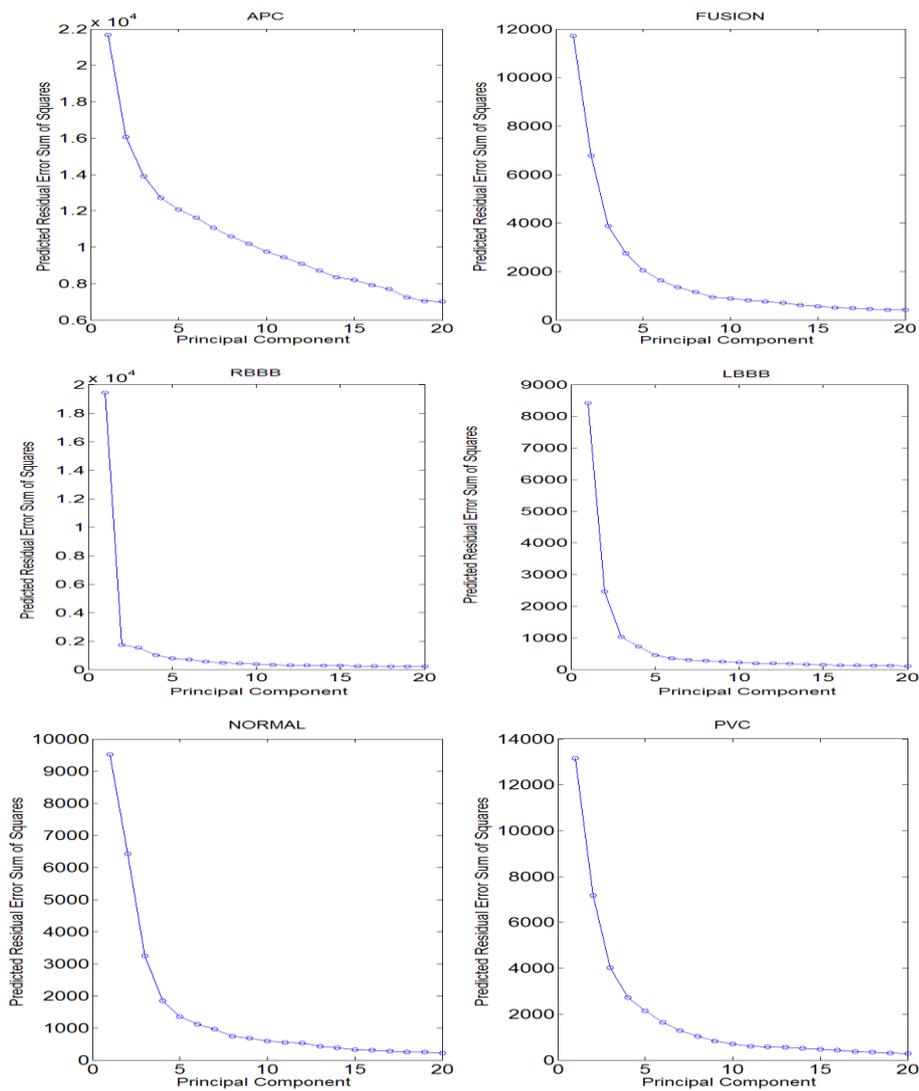
### 5.5.3.5 SIMCA Results using the precordial lead dataset for MIT-BIH Arrhythmia records

The second SIMCA study processes the MIT-BIH arrhythmia database records from precordial leads as shown in Table 5.14. There were 2400 ECG beats similar to the first study 400 Normal, 400 PVC, 400 APC, 400 FUSION, 400 RBBB and 400 LBBB beats. The cross-validation technique was used to find the optimal number of the principle components (PCs). Fig 5.28 showed the predicted residual error sum of squares (PRESS) curve for six classes using cross validation. The average predicted residual error sum of squares reached a plateau at around 20 PCs when classifying some of the classes. There was no further noticeable improvement by using more PCs. Nevertheless, this is a different result from that inferred for the limb lead study which showed that conclusions are dataset specific. In addition, Table 5.21 provides the model accuracy as a function of the different number of PCs used as input to the SIMCA classifier. As can be seen the best accuracy was obtained using 20 PCs with

93.33% classification accuracy. It was therefore shown that the dimension of each ECG beat vector can be reduced from 256 elements to 20 elements using PCA.

PCs number	Model accuracy
10	97.25
15	97.31
20	97.78
30	97.53

**Table 5.21: Classification accuracy rate for different number of PC coefficients presented to the classifier using ECG beats obtained from precordial lead (V1).**



**Figure 5.28: predicted residual error sum of squares (PRESS) curve for the six beat type classes using ECG beats obtained from the precordial lead.**

Table 5.22 illustrates the performance of the classifier in terms of the previously defined performance metrics. The proposed method misidentified 43 NORMAL beats, 9 PVC beats, 14 FUSION, and 14 LBBB beats, while all APC and RBBB beats were correctly classified. There were 13 FUSION and 6 LBBB beats that were classified as PVC beats, while 5 PVC beats were estimated as APC beats and three as FUSION beats. A total of 186 of 200 LBBB and the 200 FUSION beats were correctly classified.

Annotation	Output result					
	NORMAL	PVC	APC	FUSION	RBBB	LBBB
<b>NORMAL</b>	157	0	33	10	0	0
<b>PVC</b>	0	191	5	3	0	1
<b>APC</b>	0	0	200	0	0	0
<b>FUSION</b>	0	13	1	186	0	0
<b>RBBB</b>	0	0	0	0	200	0
<b>LBBB</b>	0	6	2	6	0	186

**Table 5.22: Confusion matrix of the ECG signal taken from the precordial lead using SIMCA classification and 20 PCs.**

Moreover, the performance of the classification for the test dataset in terms of SE, SP, and PP for SIMCA are illustrated in Tables 5.23. Six beat types can be classified with average accuracy of 97.78% using SIMCA. The best and poorest sensitivity were 100 % for RBBB and APC beats and 78.50% for NORMAL beats respectively, whereas the best and poorest SP were 100 % for RBBB and NORMAL beats and 95.90% for APC beats respectively. The sensitivities were found to be 78.50%, 95.50%, 100%, 93%, 100.0% and 93% for NORMAL, PVC, APC, FUSION, RBBB and LBBB heartbeat cases respectively. The total classification accuracy was approximately 93.33%.

Annotation	Classification performance result			
	PP (%)	ST (%)	SP (%)	ACC (%)
<b>NORMAL</b>	100	78.50	100	96.42
<b>PVC</b>	90.95	95.50	98.10	97.67
<b>APC</b>	82.99	100	95.90	96.58
<b>FUSION</b>	90.73	93	98.10	97.25
<b>RBBB</b>	100	100	100	100
<b>LBBB</b>	99.47	93	99.90	98.75

**Table 5.23: Collective result performance analysis and classification result using beats obtain from precordial lead (V1).**

Tables 5.22, 5.24, 5.25 and 5.26 illustrate the output of the confusion matrix for each beat class individually on the basis of adopting a feature reduction strategy using between 30 and 10 PCs as input to the SIMCA classifier. The feature reduction method that adjusted the length of the input

vectors for classification was repeated four times using 10, 15, 20, and 30 PCs, creating different size feature vectors. (PCs). Simulation results show that the most accurate classification that was achieved for each class was for the NORMAL beats with 79 % sensitivity, for the RBBB beats with 100%, , for the APC beats with 100% and 20 PCs. In addition, LBBB beats were classified with 100% sensitivity using 10 PCs. The best classification sensitivity values for PVC and FUSION beats were 96.50% and 94 %, respectively, using 30 PCs. Using 20 PCs, 14 of the 200 FUSION beats were estimated incorrectly, while 12 of the 200 FUSION were misclassified using 30 PCs. In addition, 9 of the 200 PCA beats were incorrectly classified using 20 PCs, whereas 7 PVC were wrongly classified using 30 PCs. Using 30 PCs illustrated an improvement in classification sensitivity of PVC and FUSION beats. Moreover, using 10 PCs illustrated an improvement in classification sensitivity and accuracy of LBBB beats. Fig 5.29 illustrates the accuracy for each beat class individually on the basis of adopting a feature reduction strategy using between 30 and 10 PCs as input to the SIMCA classifier.

Annotation	Output result					
	NORMAL	PVC	APC	FUSION	RBBB	LBBB
NORMAL	145	0	47	8	0	0
PVC	2	189	8	1	0	0
APC	0	0	200	0	0	0
FUSION	0	20	3	177	0	0
RBBB	0	0	1	0	199	0
LBBB	0	2	3	4	0	191

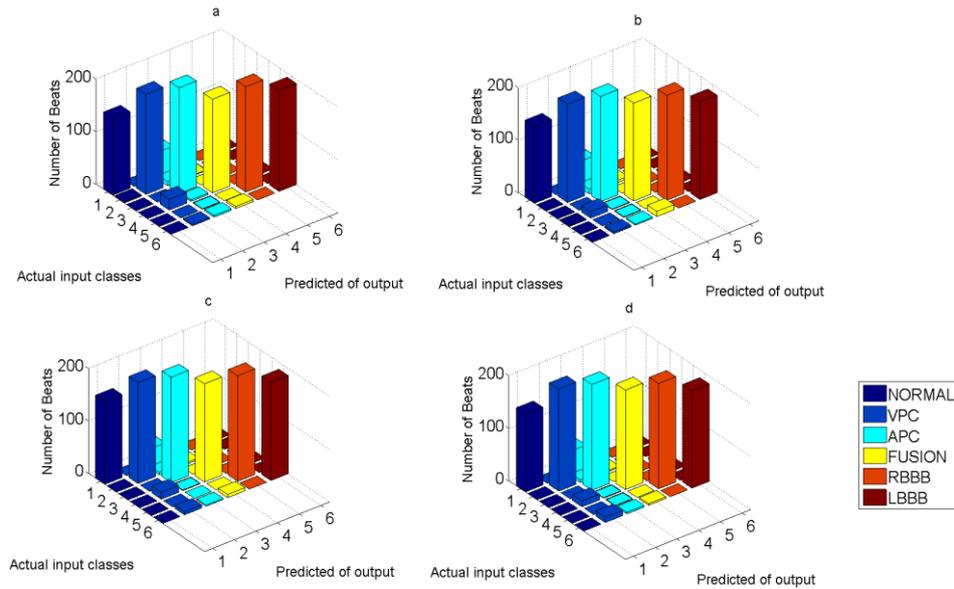
**Table 5.24: Confusion matrix of the ECG signal taken from the precordial lead using SIMCA classification and 10 PCs.**

Annotation	Output result					
	NORMAL	PVC	APC	FUSION	RBBB	LBBB
NORMAL	146	0	45	9	0	0
PVC	0	187	9	4	0	0
APC	0	0	200	0	0	0
FUSION	1	12	1	186	0	0
RBBB	0	0	1	0	199	0
LBBB	0	3	2	10	0	185

**Table 5.25: Confusion matrix of the ECG signal taken from the precordial lead using SIMCA classification and 15 PCs.**

Annotation	Output result					
	NORMAL	PVC	APC	FUSION	RBBB	LBBB
NORMAL	147	0	43	10	0	0
PVC	0	193	5	2	0	0
APC	0	0	200	0	0	0
FUSION	0	11	1	188	0	0
RBBB	0	0	1	0	199	0
LBBB	0	10	3	3	0	184

**Table 5.26: Confusion matrix of the ECG signal taken from the precordial lead using SIMCA classification and 30 PCs.**



**Figure 5. 29: the predicted of input individual ECG beat to the SIMCA classifier ECG beats taking from precordial lead (V1) using (a) 10 PCs , (b) 15 PCs, (c) 20 PCs and (d) 30 PCs.**

### 5.5.3.6 General discussion of the study classifying six classes of ECG signals

The purpose of this study was to establish if it is possible to improve multi-class classification accuracy when using a large database by developing models for each class individually so as to classify six types of heartbeats simultaneously with a high degree of overall accuracy. The first study focused in data from the limb lead (II), while the second study used ECG beats extracted from the precordial lead (V1). Classification accuracies have shown small differences from using limb leads or precordial leads beats. As can be seen the overall accuracy of using ECG beats from limb was better than precordial lead.

This result agrees with the study results of Biel *et al.*, who also found small differences in classification tests between limb leads and precordial leads data. Biel *et al.* (2011) also investigated the use of SIMCA with selected features extracted from the ECG for identifying a person in a predetermined group. The standard 12-lead ECG recorded during rest was used in that study. A SIEMENS Megacart was used to perform the ECG measurements. The information recorded dataset from this was transferred and converted to a usable matrix format ( $30 \times 12$ ). Features such as P wave onset, P wave duration, QRS wave onset, and QRS wave duration were extracted for each person to be used as inputs to the SIMCA classifier. To reduce the amount of features to 12, the correlation matrix was employed. This reduced the dimensionality of features for the input vector. SIMCA was used to classify persons and identify individuals with 95% accuracy [232]. To my knowledge the

current study presented in this thesis provides a new approach to SIMCA classification of ECG measurements. Since PC features were used, the in this thesis studies are the only ones that have been done using a SIMCA PC classifier.

The results presented show that it is possible to simultaneously classify the six ECG beats effectively. What is still needed is a comparative discussion regarding the performances of the feature reduction techniques, the different classification techniques, and the entire arrhythmia classification procedure. Results are therefore placed in a more general perspective by considering different feature reduction approaches and classifier methods presented in the literature.

In [233] an intelligent diagnosis system using an Adaptive Neuro-Fuzzy Inference System (ANFIS) model was built to differentiate among six types of ECG signals, namely normal sinus rhythm (NSR), PVC, APC, ventricular tachycardia (VT), ventricular fibrillation (VF) and supraventricular tachycardia (SVT). An Independent Component Analysis (IPCA) and power spectrum information were used to extract input feature vectors of the ANFIS classifiers. The simulation results indicate a high classification accuracy at more than 97%. Patra *et al.* (2010) investigated the classification performance of six types of ECG beats using four structures: FCM-NN, PCA-NN, FCM-ICA-NN, and FCM-PCA-NN. Fifty sample segments attributed to six ECG beat types were selected from the MIT-BIH arrhythmia database for this study. These six beat types were NORML, LBBB, RBBB, APC, PVC, and paced beats (PB). The fuzzy c-means (FCM) algorithm was used to decrease the number of segments by grouping similar segments in the training data, whereas feature extraction was performed using PCA. Using the FCM-PCA-NN structure produced a better result than other techniques was obtained with an average error rate of 0.13% [168]. Six different features of ECG signals were used to discriminate and classify the same six types of heart beat. Three of the features were calculated statistically from decomposed sub-band signals that were obtained using a discrete wavelet transform, while two features were taken from the AC power and the instantaneous RR interval of the original signal. PNN was utilised to classify six types of heart beats using just 11 features for each type as input vectors for the classification. The performance of classification using this feature set from a wavelet transform as an input to the PNN classifier achieved an accuracy of 84.35% [234]. In 2008, Korürek and Nizam proposed a new arrhythmia clustering technique based on Ant Colony Optimisation (ACO) and neural networks. This technique was tested using ECG beats

extracted from the MIT-BIH database to classify six different arrhythmia types, namely normal sinus rhythm, PVC, APC, RBBB, ventricular fusion and fusion beats. The sensitivity simulation results were produced using two parameters together with ACO and neural networks, these had values of 93.05 and 94.40 respectively [235]. In [17] three different classification techniques, support vector machine (SVM), multilayer perceptrons (MLP) and the fuzzy inference system (FIS) with reduced features using of LDA and PCA were adopted to classify six types of arrhythmia beats: NSR, APC, SVT, PVC, VT and VF. Accuracy of classification of NSR, APC, SVT, PVC, VT and VF were 99.307%, 99.274%, 99.854%, 98.344%, 99.441% and 99.883%, respectively. Asl *et al.* (2008) utilised the generalised discriminant analysis (GDA) feature reduction method to select five features from the 15 different features extracted from the input HRV signal by means of linear and nonlinear methods. Subsequently, these five characteristic features and an SVM classifier were combined with an one-against-one (OAO) and an one-against-all (OAA) strategy. These were then used to discriminate between six different types of arrhythmias with average accuracy of 98.90 and 99.16, respectively [104]. Using a local fractal dimensions (LFD) algorithm, six types of ECG beats were classified using an estimation of the LFD of ECG signals as features and the nearest neighbour classifier approach. To estimate the LFD, two methods were adopted, namely the power spectral density-based fractal dimension estimator (PSDFE) and the variance-based fractal dimension estimator (VFE). Normal, LBBB, RBBB, APC, VPC and PB beats can be categorised with 88.64% average sensitivity, with an individual performance of 93.15% for normal beats and 91.07% for VPC arrhythmia [236]. In another study, the wavelet transform and particle swarm optimisation techniques were proposed to extract features from six types of ECG beats. The classification performance of these six types of ECG beats (normal, LBBB, RBBB, APC, VPC and PB) was performed using SVM, this study showed an overall accuracy ratio of 88.84% [237]. Doğan and Korürek suggested a new clustering method based on a kernelised fuzzy c-means (KFCM) algorithm and a recently proposed ant-based optimisation algorithm, the hybrid ant colony optimisation for continuous domains (HACO), to classify six types of ECG beats: NORMAL, PVC, FUSION, APC, RBBB and Fusion of Paced and Normal Beat (FUSIONP). The average sensitivity and specificity values were revealed to be 93.76 and 98.76 respectively when using the KFCM-HACO algorithm [238]. In addition, in another study six types of beats, including a NORMAL, PVC, FUSION, APC, RBBB and FUSIONP beats were classified using particle swarm optimisation (PSO) and a radial basis function neural network (RBFNN) with an

average sensitivity of 95.24 and a specificity of 98.9 [113]. In another study, the same authors proposed a switchable scheme based on IPCA to determine six different types of ECGs (NORMAL, LBBB, RBBB, APB, PVC, and PB) with an accuracy rate of over 99% [99]. In addition, the overall performance of the PCA-LDA-PNN scheme was better than the aforementioned ECG arrhythmia classification algorithm. The proposed scheme in this thesis therefore seems to show the best overall result for the simultaneous classification of six types of ECG arrhythmia. The SIMCA classifier can outperform the abovementioned ECG arrhythmia classification method because it can significantly reduce the number of misclassifications. According to Table 5.17 the average sensitivity, specificity, and accuracy were 99.06%, 99.61% and 97.98%, respectively by using the proposed ECG arrhythmia classification method suggested in this study. The proposed scheme can provide better classification accuracy compared to the other schemes found in the literature and maybe considered as an effective tool for diagnosing heart disease.

#### **5.5.4 Study four: Classification structure and results using SIMCA and MSVM classifiers**

##### **5.5.4.1 ECG beat acquisition and pre-processing**

Four types of beats including, NORMAL, PVC, APC and RBBB beats were extracted from the St Petersburg INCART 12-lead Arrhythmia (incartdb) database and used for this study. A number of beats that showed simultaneously characteristics from all the four type beat types mentioned were selected, these were associated with recordings from limb leads II and III from the following four patient records: I17, I20, I22 and I71. In total, 1800 heartbeat signals were extracted from this database. These beats were used for classifier training as well as for evaluating MSVM and SIMCA classifier performance.

The features selected in the current study focus on the segment associated with the P, QRS complex and T wave. Firstly, the R peak location is identified using the annotation file; then R (peak) location is used as a reference to detect peaks in other waves (such as P and T). Each beat within a segment includes 159 samples before and 160 samples after the R event; the window function, therefore, spans across 320 points, these data points are approximately equally spaced every 0.889 s as discussed in [239]. As stated earlier, the extracted ECG beats had their baseline removed before PCA feature extraction and formulation of the input vector to the classifier.

### **5.5.4.2 Feature extraction**

In the present study, as a starting point between 10 and 30 PCs were selected for each beat as input features for classification. This number was then adjusted through the procedure of determining the optimal number of PCs. The best number of PCs for each beat class was determined through the cross-validation process. This way, the dimension of each ECG beat sample was reduced from 320 to 30 or fewer.

### **5.5.4.3 NORMAL, PVC, APC and RBBB beat classification results from the MSVM and SIMCA study using two leads (II, III).**

The MSVM and SIMCA classifiers were used to distinguish between the four ECG arrhythmia types on the basis of the pruned PC coefficients forming feature vectors at the input stage of each classifier. MATLAB 2010a software routines were used to train and validate both the MSVM and SIMCA classifier, with LIBSVM used to train and validate MSVM. For each experiment 1800 beats from the four ECG beat types were selected. Each dataset included 600 NORMAL segments, 400 PVC segments, 400 APC segments, and 400 RBBB segments. The normalised ECG beat samples were randomly divided into two equally sized groups consisting of 900 beats each, for training and testing respectively. Once the optimal dimensions for each beat class were obtained, the classification performance of both the SIMCA and MSVM classifiers was compared after presenting all four types of ECG arrhythmias simultaneously.

#### **5.5.4.3.a MSVM classifier structure**

LIBSVM software routines were used to train and validate the MSVM model. Sequential minimal optimization (SMO) [188] was used to compute hyper-plane parameters and threshold values for MSVM during the training phase. The MSVM and kernel parameter values were obtained using 5-fold cross validation. Firstly, 5-fold cross validation was adopted for adjusting the kernel and SVM parameter values. Then, SMO was used to train the MSVM and compute the corresponding optimal hyper-plane parameters. The optimal hyper-plane parameters were identified on the basis of the training data set and class label. Finally, on the basis of the training results, the test datasets were imported to the MSVM classifiers to perform unknown beat classification. The optimal valued for ( $\alpha_i$ ) and threshold ( $b$ ) were adopted and applied to the test datasets in order to measure the classification performance for MSVM classification.

#### 5.5.4.3.b SIMCA classifier structure

As mentioned earlier, the advantage of SIMCA is that it is based on creating individual models for each class and finding similarities between test objects and class models. The first step in SIMCA modelling is to build a model for each class using an optimal number of PCs (following the cross-validation procedure) and creating a training dataset. Since the average predicted residual error (sum of squares) reached a plateau at around 30 PCs in the majority of classes and no noticeable improvement could be observed using more PCs, this was taken as the optimal vector size. Our aim was not to overwhelm the classifier as this would compromise its generalization capability. After deciding on the number of PCs used, individual models for each beat class was created. These models were tested with the test dataset samples to measure the classification performance attained.

#### 5.5.4.3.c Simulation results

The classification performance was evaluated after presenting 900 beats associated with each of the four ECG beat types. To evaluate the performance of the proposed arrhythmia classification algorithm, the commonly adopted metrics of sensitivity (SE), specificity (SP), positive predictivity (PP) and accuracy (ACC) were used. Moreover, the classification results for each ECG beat type are displayed using confusion matrices, which illustrate the success of the classification process in terms of classified or misclassified results, these are shown in Tables 5.27 and 5.28

Table 5.27 presents the confusion matrix of the classification results using SIMCA, whereas Table 5.28 those from MSVM. It can be seen that all RBBB are correctly classified using the MSVM algorithms. With MSVM, 121 N beats and 7 PVC beats were misclassified as an APC beat, while with SIMCA only one PVC beat and 12 N beats were misclassified as an APC beat. Six RBBB beats and 11 APC beats were classified as PVC beats using SIMCA. In contrast, when using MSVM, 5 APC beats were misclassified as PVC beats.

In addition, Table 5.29 summaries the performance result of the classification process using the four common metrics: *ST*, *SP*, *PP* and *ACC*. The proposed algorithm achieved classification accuracies up to 98.33%, whereas MSVM achieved up to 76.83 % accuracy. Moreover, the sensitivities were found to be 96%, 99.5%, 94.5%, and 100.0% for N, PVC, APC and RBBB cases, respectively using SIMCA. The sensitivities were 59.3%, 73.5. %, 79 %, and 100% for N, PVC, APC and RBBB beat cases respectively using MSVM. Using

SIMCA, there was an improvement in the sensitivity for all classes and significantly enhanced the overall classification accuracy. The SIMCA classifier combined with PCA demonstrated excellent performances in discriminating four beats types with average accuracy, sensitivity, positive predictivity and specificity of 98.33%, 96.75%, 96.4% and 98.9% respectively.

Annotation	SIMCA output result			
	NORMAL	PVC	APC	RBBB
NORMAL	288	0	12	0
PVC	0	199	1	0
APC	0	11	189	0
RBBB	0	6	0	194

**Table 5.27: Confusion matrix using SIMCA and the St Petersburg INCART 12-lead Arrhythmia Database (incartdb).**

Annotation	MSVM output result			
	NORMAL	PVC	APC	RBBB
NORMAL	178	0	121	0
PVC	13	147	7	33
APC	37	5	158	0
RBBB	0	200	0	0

**Table 5.28: Confusion matrix using MSVM and the St Petersburg INCART 12-lead Arrhythmia Database (incartdb).**

Annotation	Classification performance result							
	SIMCA				MSVM			
	SE	SP	PP	ACC	SE	SP	PP	ACC
NORMAL	96	100	100	98.6	59.3	91.6	78	80.9
PVC	99.5	97.5	92.1	98	73.5	70.5	41.6	71.2
APC	94.5	98.1	93.5	97.3	79	81.7	55.2	81.1
RBBB	97	100	100	99.3	100	95.2	100	74.1

**Table 5.29: Collective PP, SE, ACC and SP results using MSVM and SIMCA classifiers for a 2-lead study using the St Petersburg INCART 12-lead Arrhythmia Database (incartdb).**

#### 5.5.4.3.d Discussion

The proposed SIMCA algorithm achieved classification accuracies up to 98.33%, whereas MSVM up to 76.83% accuracy. Additionally, these results have been compared with other ECG beat classification systems such as PCA with SVM [213], SVM with wavelet transform and Fourier Transform [211] and DAGSVM. Table 5.30 compares the recognition rate of some of these methods from the current literature with the proposed scheme. It is clear that the use of SIMCA provides better classification results in terms of recognition rate.

Feature reduction method	Classification method	Overall Accuracy (%)	References
DWT	SVM	94	[211]
DCT	SVM	96.5	[211]
Amplitude value	SVM	94.2	[211]
Genetic Algorithm	SVM	93.46	[213]
PCA	SVM	80.00	[213]
Lyapunov exponents	RNN	94.72	[6]
SVD	DAGSVM	97.71	[224]
DFT	CSVM	94	[28]
PCA	MSVM	76.83	Current work
PCA	SIMCA	98.33	Current work

**Table 5.30 Comparison study four result with other classifiers.**

## 5.6 Summary

In this study, beat classification was performed using MSVM, CSVM and SIMCA classifiers. All results were obtained using MATLAB 2012a software routines for training and validation. Extracted feature in both time and frequency domain were considered. In order to improve the classification accuracy of the classifier, the dimensional features were reduced using several feature reduction methods. The SMO was used to compute hyper-plane parameters and threshold values for both MSVM and CSVM during the training phase. Using CSVM it was possible to successfully classify four types of ECG beat signals simultaneously. The algorithm indicated higher performance than MSVM classifiers. Results illustrate that the proposed beat classifier is very reliable, and that it may be adopted for automatic detection of arrhythmia conditions and classification. Using CSVM, a 4-class problem can be classified rapidly by decomposing it into two distinct SVM tasks. The overall accuracy using ECG beats extracted from limb lead and precordial leads were 97% and 100% respectively. Moreover, the present research confirmed that the use of selected number of Fourier coefficients to approximate the ECG beat signal and compress the input features to the classifier can lead to high classification accuracies and improve the generalization ability of the classifier. To our knowledge this is the first time that a multi class algorithms such as MSVM and CSVM were employed in ECG beats extracted from the European ST-T Database and the first time that a CSVM classifier was used with the other database as well.

The SIMCA classifier was used to differentiate between six types of ECG beats. The PC feature vectors around each R peak of ECG beat were obtained from the MIT-BIH database and the St Petersburg INCART 12-lead Arrhythmia Database (incartdb). The optimal number of features was presented at the inputs of the SIMCA classifier to determine the six different types of ECG arrhythmias. The average classification accuracy of the proposed scheme was 98.61% and 97.78% using limb and precordial lead datasets, respectively. In addition, MSVM and SIMCA classifiers were used to differentiate between four different types of ECG beats obtained from the St Petersburg INCART 12-lead Arrhythmia Database (incartdb). The average classification accuracy of the MSVM and SIMCA classifiers were 76.83% and 98.33% respectively. It is concluded that SIMCA could be an effective tool to diagnose arrhythmia conditions using ECG signals. The effectiveness of the proposed scheme for ECG arrhythmia classification was proven by successfully classifying the six types of ECG with a high average accuracy ratio. It is speculated that the

proposed scheme may be able to identify more than six beat types with good classification accuracy. The following chapter summarizes the findings of this study, and provides some recommendations for further work.

## Chapter 6 Conclusions and future work

### 6.1 Conclusions

This study has introduced novel classification techniques, which can efficiently detect and classify cardiac events. The work adopts CSVM and SIMCA classifier algorithms to solve multiclass problems in ECG beat classification. These were used on records selected from three databases: the MIT-BIH arrhythmia database, the European ST-T Database and the St Petersburg INCART 12-lead Arrhythmia Database (incartdb). The proposed methodology provides a useful way for analysing and classifying arrhythmias conditions based on supervised classification approaches. The CSVM algorithm demonstrated excellent performance, mapping the extracted ECG features to a higher dimensional space through the use of two complex hyper-planes. This has been developed as an extension of the well-known SVM classification algorithm. The performance of the SIMCA classifier was based on creating individual models for each class and finding similarities between test objects and class models. The implementation of both CSVM and SIMCA focused on classifying heart beats into the right categories. The general approach adopted was composed of three-stages: pre-processing and beat extraction, feature reduction and finally classification. The importance of feature selection for ECG classification has been the subject to considerable discussion, this work confirmed its necessity and merits.

Several studies have been presented and discussed in order to measure the effectiveness of the proposed methods. Using CSVM, a 4-class classification problem can be solved rapidly by decomposing it into two distinct SVM classification tasks. The first study was focussed on using a binary SVM classifier to differentiate between normal and abnormal beats that were extracted from the different lead. A high classification accuracy of 100% was obtained from using this technique. For CSVM, the four types of beats were obtained from the MIT-BIH arrhythmia database first. The ECG beats were obtained using a window function composed of 300 samples around the R-peak. Furthermore, the work proposed a methodology that combines the use of filtered coefficients from a DFT in conjunction with CSVM to improve the generalization ability of the classifier. In order to estimate the best number of Fourier coefficients, between 100 and 30 Fourier coefficients were used to

reconstruct individual ECG beats using the most prominent features in the frequency domain. In the associated DFT, the complex values in the signals were retained as vectors. These vectors were used as the input feature vectors for the training and testing of the CSVM. For the classifier implementation, after feature selection was performed, the datasets were divided into two groups for training and testing purposes. The optimal hyper-plane parameters were identified on the basis of the training dataset and class label. The introduction of the CSVM framework as an extension to normal SVM was made on the basis that our intention was to preserve information in both amplitude as well as phase in the ECG signals. Accuracies between 83% and 97% are obtained for MSVM and CSVM classification respectively. Using the CSVM algorithm indicated a significant improvement in multi-class classification over MSVM. Using the CSVM classifier algorithm indicated higher performance than MSVM classifiers. In another study, using the CSVM algorithm with the ECG beats that were extracted from four precordial lead provided high classification accuracies with 0.8% error rate when trained with fifty DFT coefficients, and 1.75% error rate when trained with 100 DFT coefficients. The proposed algorithm demonstrated classification accuracies of 99%, 100%, 98%, and 99% in datasets obtained from one, two, three, and four leads respectively. Moreover, using the ECG beat from two correlated leads showed a significant improvement in the classification accuracy, with an average accuracy of 100%.

In another study, the ECG beats were obtained from the European ST-T Database with 212 samples around the R-peak. In the feature selection step, 100 Fourier coefficients were selected for reconstructing individual ECG beats, these formed the input vector to the classifier. On the basis of the training results, test data were imported to the CSVM and MSVM classifiers to perform unknown beat classification. The proposed algorithm confirmed classification accuracies of 94%, whereas multi-class SVM achieved accuracies up to 86%.

This study also evaluated the use of the SIMCA classifier differentiate between six types of ECG beats. It was demonstrated that the SIMCA classifier can be used to simultaneously differentiate between the six types of ECG beats even if windows had different dimensions. The 256-dimensional feature vectors around each R peak was obtained from the MIT-BIH database. In order to improve the classification accuracy of the classifier, the 256-dimensional features were reduced to between 10 and 30 features using PCA. The optimal number of features and the RR time interval were treated as inputs to the SIMCA classifier which was assigned to simultaneously to determine six different types of ECG

arrhythmias. The average classification accuracy of the proposed scheme was 98.61% and 97.78% using limb and precordial lead datasets, respectively. The effectiveness of the proposed scheme for ECG arrhythmia classification was confirmed after successfully classifying six types of ECG with a high accuracy ratio.

The final study evaluated the integration of MSVM and SIMCA classifiers with PCA extracted features to differentiate and correctly classify four different types of ECG beats obtained from the INCART 12-lead Arrhythmia Database (incartdb). All the different ECG beat types were presented to the classifier simultaneously. The 230 point feature vectors around each R peak of each ECG signal were reduced to the optimal number of features following PCA decomposition and extraction of PCs. The average classification accuracy of the MSVM and SIMCA classifiers were 76.83% and 98.33% respectively. The SIMCA classifier algorithm was able to determine four beat types that were simultaneously presented to it with a better classification accuracy than the MSVM classifier.

The use of the CSVM algorithm indicated higher performance than did the MSVM classifiers. The CSVM classification algorithm was able to classify four classes of ECG arrhythmias with an accuracy rate between 94% and 100%. Through the feature extraction phase, it was possible to reduce the sample set on each heartbeat for example from 301 samples to 50 samples while maintaining a high accuracy rate. These numbers of features were found to give good classification accuracy with good specificity and sensitivity rates as well.

In addition, a reduction of samples could reduce the memory size and processing time of classification without compromising on accuracy. The outcome of the present research confirmed that the use of a selected number of Fourier coefficients to approximate the ECG beat signal and further compress the input features to the classifier can lead to high classification accuracies and improve the generalization ability of the CSVM classifier.

Furthermore, using the SIMCA classifier, it was possible to successfully simultaneously classify six types of ECG beat signals with different dimensions. The SIMCA classifier was employed to create models for each class individually. It is believed that using the SIMCA scheme proposed, in the future could enable the determination of more than six beat types with good classification accuracy. It was concluded that both CSVM and SIMCA could be effective tools to diagnose arrhythmia conditions. To our knowledge, this is the first

time that a CSVM and SIMCA algorithm were employed in ECG studies to simultaneously classify four and six types of ECG beats respectively. The work demonstrated the effectiveness of the proposed scheme for ECG arrhythmia classification by achieving an overall high accuracy ratio. Using the CSVM algorithm indicated a significant improvement in multi-class classification over MSVM. The results also showed a high specificity and sensitivity when using CSVM for ECG signals classification.

## **6.2 Limitation of the performed studies**

It should be borne in mind that the performed studies have the following limitations:

- The scope of the SIMCA classifier is limited to the development of an ECG classification algorithm based on PCA using time-domain and not frequency-domain features.
- Using the CSVM classifier, the major limitation is that only four types of ECG beats can be classified simultaneously. Furthermore, a larger number of ECG recordings classified would have given us a more robust classification result than the current one.
- Due to limited number of studies performed, the algorithms in this work may require further validation using both time domain and frequency domain features as input vectors to the classifier.
- The small sample size in some classes did not permit further systematic investigations using SIMCA with the European ST-T Database. One of the drawbacks of the SIMCA classifier is that it requires a large dataset for training and testing.
- The European ST-T Database and St. Petersburg INCART12-lead Arrhythmia Database (incartdb) do not include different types of ECG beats like the MIT-BIH Arrhythmia Database does, so an inter-comparison of the developed algorithms across all databases was not possible.
- Due to the absence of 12 ECG-lead records in popular ECG databases, such as the European ST-T Database and MIT-BIH Arrhythmia Database, in some of our studies we were restricted to using only up to four leads.
- The study did not have access to the newly developed multi-lead recordings based on the MEDTRONIC vest. It would be useful to apply the proposed algorithms to those datasets.

## **6.3 Future work using CSVM and SIMCA**

The proposed approach is to be extended using adaptive wavelet where a different wavelet function is derived at each decomposing level to increase parsimony of the input

vector. For feature selection wavelet pre-processing will be used to further compress the input to the classifier by generating wavelets on the basis of higher order moments. It proves convenient to write the transform in matrix form as discussed in chapter 4. A test waveform was decomposed to qualitatively describe the new parametrization process as would be implemented. This decomposition was contrasted to that of a standard db6 parametrization. It was originally intended to further explore this approach as a feature reduction method before presenting ECG signals from the MIT-BIH database to CSVM classifiers, but such study was postponed due to time constraints for submission of the thesis.

Multidimensional SVM using Clifford Algebras should also be considered to account for multi-lead signal analysis. Some of the future work should therefore concentrate on the further development of ECG signal pre-processing using adaptive wavelet algorithms, as combined with the classification with Clifford SVM techniques. The aims of using Clifford Algebras SVM to adopt CSVM algorithms for heterogeneous (multi-lead or multi-sensor) signals in complex spaces (to account for signal amplitude and phase information separately or account for time-frequently features in the ECG signal on the basis of wavelet analysis) as well as examine the possibility of adopting multi-dimensional Clifford algebra frameworks that will cross-correlate information at the input space of the classifier are important directions that are likely to make significant impact in the general area of ECG signal analysis. The potential sensitivity of the technique might give clues on the impact of diet and lifestyle directly to the overall health state of a heart.

Additionally, different wavelet decomposition schemes of ECG signals assuming different levels of approximation and detail for each wavelet family will be considered and their suitability as inputs to the CSVM classifier is going to be investigated. The efficiency in parametrizing the wavelet coefficients using adaptive structures from the perspective of improving parsimony in each decomposition step and improving the reliability of the classification task would have to be investigated further. The proposed approach intends to optimize the number of wavelet coefficients as well as a number of decomposition levels presented to the classifier. In addition, the proposed algorithm may be extended to higher dimensional spaces associated with division two algebras for the simultaneous classification of four or eight classes of inputs so it represents a generic methodology of relevance to future multi-lead recordings as well as to other heterogeneous emergent ECG monitoring modalities.

Looking at the problem from the perspective of providing in the future complete system solution one should note is that the mobile-phone industry has been rapidly developing recently. There are multiple features that have been integrated into the mobile phone nowadays. Current smartphone models tend to integrated several features of other biosensor and have been influenced by recent developments in the television industry. For example, smartphones are also used to view videos, as well as access the internet; such devices have been equipped with resolution capabilities that match those of TV devices. Such innovations may not only be used for entertainment purposes but are also likely to be used for other functions such as medical- and health-related features. Furthermore, more mobile-phone manufacturers are also looking for ways of improving their current product offerings through extensive research and development. The use of new mobile algorithms for ECG classification and the diagnosis of some long-term conditions such as heart diseases are a natural direction for further investigation. The current study predicts that the use of smartphones and a wireless communication networks for monitoring patients at home or on the go over a prolonged period of time (holder function) for the purpose of providing early diagnosis of heart conditions is very likely to further proliferate. The ability of feature reduction of the ECG signal has additional advantages in mobile platforms such as low memory.

As a final thought, it is also believed that the current modality of mapping heartbeats using electrodes may also change in the future through the use of non-contact magnetic sensors though the development of nano-SQUID (superconducting quantum interference devices). Such signals would not be spatially confined to the location of the electrodes used in current practice, giving rise to new opportunities for heart monitoring. Such sensors would provide a better picture of any abnormal electric wave propagation in the human body. A further area of research that can directly benefit from the current work is that of deep AI as applied to ECG analysis. In this context the aim is to use multi-layered neural networks with different levels of feature abstraction at each layer that needs to be trained. A geometric neuron based on Clifford algebras would be the most versatile structure to be adopted in this context.

## **6.4 Trends in mobile computing: using mobile devices and wearable sensors to analyse ECG signals**

As the development of next-generation wireless technology such as GPRS, EDGE and 3G using mobile devices has been increased in the health sector, the mobile device has become a possible tool for telemedicine because some medical information can be transferred to or from patients through their mobile phone. A good illustration of this is the transmission of an electrocardiogram ECG signal to a mobile by using Bluetooth and Multimedia Messaging Service (MMS). Nowadays, the majority of smart phones are able to send and receive MMS which means that special application software is not needed to display the ECG. It is anticipated that in the near future patients will be able to import electrocardiogram ECG signals as a bitmap image through their mobile phone when at home or outside the hospital [240]. One of the advantage through telemedicine is the possibility for long-term monitoring of patients with cardiovascular diseases at home, this is made possible through the use of portable ECG telemonitoring systems. Such system is normally divided into two modes: store-and-forward mode (patient data is available at a later time); and real-time mode (the patient's data is immediately available at the server end after acquisition). A patient-worn unit called Holter, can continuously provide information of a patient's health through MMS on GPRS signals or through direct connection to internet using Wi-fi [241]. Due to the considerable cost of health care, remote monitoring of medical conditions is rapidly becoming an alternative for monitoring patient health states. The main aim for such future developments is to focus on the complete architecture of a mobile/wireless communication in conjunction with effective implementation of the algorithms, this should allow doctors to have access to the results from the ECG data analysis immediately [242].

With the advancement of technology comes the breakthrough possibility of effectively safeguarding human lives. Because of the technological advances in the field of mobile telecommunications networks, telemonitoring is widely used by more than half of the US home care agencies. Not only it is cheaper, it also gives the patients the liberty to stay at home with their family. Development in mobile devices such as smart phones and in programming platforms such as J2ME and Android has led to a recent proliferation of mobile-based telemonitoring system. ECG signal monitoring and analysis procedures use the computing capability of mobile devices. This system is very convenient as the mobile phone can be handheld by both patients, or medical experts. In an experiment conducted by Fang's team

[243], it was proven that using a J2ME platform, a mobile device can be programmed so as to be used for reliable automatic bio-signal collection, noise identification and reduction, and message packet creation and transmission. Biosignal transmission can be made possible through SMS and MMS. ECG data compression and decompression, and pathological pattern detection and encryption are some of the tasks that can be performed by a mobile phone that is programmed by a JavaTM based software [243]. Some of these developments are also of interest to NASA for astronaut health monitoring.

The data flow lets the patient to be always in contact with his medical doctor for monitoring his medical conditions. For instance, the results and analysis of an ECG recording, after being checked by the doctor, can be sent to the patient for data logging and storage through GSM/GPRS communications. Connected Limited Device Configuration (CLDC) and Mobile Information Device Profile (MIDP) may be also used for application development on mobile devices. Different sub-programs which can store and enter ECG data in the database must be present in the PC for this to be possible [242]. This system provides economical and easily realizable solution for telemedicine in which medical monitoring can be done to the patient anytime and anywhere.

In 2008, Young Chung *et al* pioneered mobile applications capable of monitoring ECG waves by detecting the QRS complex and determining the normality of the vital signs. They have built a system that performs mobile-phone local vital sign data analysis and transmits data over a wireless sensor network. The patterns of signals can be identified by doing simple data analysis and then immediately transmitting these signals to a hospital server for diagnosis. The PanTompkins QRS detection algorithm was used for the QRS Detection, with 99.3% of the QRS complexes successfully detected correctly, in studies based on signals from the MIT-BIH Arrhythmia Database.

In the same year, another mobile personal electrocardiogram monitoring system integrated with a transmitter based on MMS technology was also developed. Tahat (2008) discussed the use of MMS for transferring a patient's ECG signal and body temperature. The ECG signals were taken from the three ECG electrodes left arm (LAR), right arm (RAR), and right leg (RL) these were fed into the inputs of the designed instrumentation amplifier and conditioning circuit. The application software running and displaying the received ECG signal and temperature via Bluetooth had a capture button that could be used for saving the display as a bitmap image that could also be sent as an MMS message [240].

To conclude, telemedicine through mobile sensor integration and signal processing has been widely considered to be part of the inevitable future of the modern practice of medicine. It could lead to a new approach of practice for patients, helping them to monitor their health at home, providing early diagnosis, treatment, and increased convenience. New communications technologies, such as GPRS, EDGE, 3G, and WiMax, can be used to provide much higher data-transmission speeds (rates) compared with the basic 2G GSM mobile-phone system. There should be, therefore, considerable scope to further adopt the algorithms discussed in this thesis to such mobile platforms.

### **6.4.1 Future work on multi-lead analysis using Clifford algebra and Clifford Support Vector Machines.**

#### **6.4.1.1 Future work in ECG Multi-lead analysis**

Multilead ECG analysis is what is being implemented and developed in several hospitals nowadays. A number of approaches have been established in order to reduce and then eliminate the several disadvantages of offline ECG analysis, such as the inaccurate quality control upon data acquisition and therefore, lack of quality of the resulting data.

In an attempt to improve the quality of data obtained in ECG analysis, multilead estimation of T-wave alternation in ECG was explored in the paper of Monasterio *et. al.*, in 2008. In this study, it was proposed that a combination of PCA and the generalized likelihood ratio test (GLRT) would be a better way to estimate the T-wave alternation in ECG [244]. The performance of this novel combination of extracted features, was assessed and compared with that of the single-lead scheme. Moreover, the accuracy of the estimation was calculated independently. It was found that a multi-lead scheme is better than a single-lead scheme because it has a lower variance. The low variance would mean that the final relative error in the data is also low. The multilead scheme could detect T-wave alternans (TWA) with an SNR of 30 dB lower and estimate TWA with an SNR of 25 dB lower than the single-lead scheme.

PCA was also applied for the removal of noise of Multichannel ECG signals. This is for the purpose of improving the quality control of data provided by the ECG. Sharma *et al.*, applied the PCA to wavelet scale multivariate data matrices to remove signal noise. The wavelet scale multivariate data matrices are formed from the collection of wavelet coefficients of all ECG channels. In order to enhance the PCA, the signal-to-noise ratio was

computed, and percentage root mean square difference (PRD) and wavelet energy based diagnostic distortion (WEDD) were applied to evaluate the signal distortion metrics. It was found that using the given strategy, the SNR was lowered with better denoising effect [245].

This was followed by another approach, in which a novel multi-lead-based automatic strategy was performed to delineate the ECG boundaries with respect to the QRS and T-wave boundaries [246]. The single-lead and multi-lead (ML) were both subjected to delineation and a generalized algorithm was used for multi-lead boundary location. In order to validate the automatic delineation strategy, this was used over files from the available ECG databases. The results of this study showed that the automatic ML approach can be applied to multiple leads. The given strategy provided more robust and accurate boundary locations using the delineation system.

In another strategy used by Homaeinezhad *et al.*, in 2012 sequential particle SVM was used in categorizing the sequential heart arrhythmia. In this approach, baseline wander was removed and the computation of scaled multi-lead ECG signals was done. A classification algorithm was also applied for accuracy. Because of the sequential particle support vector machine that was adopted, heart arrhythmia classification performance was improved. The specificity and accuracy, as well as the sensitivity and productivity also increased [247].

In order to obtain an improved quality control, Guillemard *et al.*, used Clifford Algebras and Dimensionality Reduction to separate and classify signals [248]. The authors projected the data to different directions across an orthonormal basis to filter and organize it. The results showed that these approaches can be used in organizing, enhancing and classifying data that are too hard to decipher.

#### **6.4.1.2 Clifford algebra and Clifford Support Vector Machines for multi-lead ECGs**

The main goal in Bayro-Corrochano and Arana-Daniel's study: 'Clifford Support Vector Machines for Classification, Regression, and Recurrence,' was to generalize the real-valued SVMs to the hyper complex or Clifford-valued SVMs, and develop the concept of multiple input multiple output (MIMO) CSVMs, and to utilize CSVMs as classifiers, regressors, and recurrent systems. Their research focused more on the Recurrent Clifford Support Vector Machines (CASVMs) and is of much relevance to the work presented in the previous chapters as it can efficiently tackle multi-lead datasets [231]. This is an important direction for future research in ECG classifiers. Implementations can be envisaged using

either multi-dimensional kernel functions based on SVM approaches or using neural networks and geometric neurons.

## References

- [1] A. L. Goldberger, Z. D. Goldberger, and A. Shvilkin, *Goldberger's Clinical Electrocardiography A Simplified Approach*, 8th ed. Philadelphia: Saunders: Elsevier, 2013.
- [2] N. Maglaveras, T. Stamkopoulos, K. Diamantaras, C. Pappas, and M. Strintzis, "ECG pattern recognition and classification using non-linear transformations and neural networks: a review," *International Journal of Medical Informatics*, vol. 52, no. 1–3, pp. 191–208, 1998.
- [3] G. G. Herrero, A. Gotchev, I. Christov, and K. Egiazarian, "Feature extraction for heartbeat classification using independent component analysis and matching pursuits," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05)*, 2005, pp. iv–725–iv–728.
- [4] G. S. Nayak and D. Nayak, "Classification of ECG Signals using ANN with Resilient Back Propagation Algorithm," *International Journal of Computer Applications*, vol. 54, no. 6, pp. 20–23, 2012.
- [5] A. Khazaee and A. Ebrahimzadeh, "Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features," *Biomedical Signal Processing and Control*, vol. 5, no. 4, pp. 252–263, 2010.
- [6] E. D. Übeyli, "Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1192–1199, 2010.
- [7] E. D. Übeyli, "Combining recurrent neural networks with eigenvector methods for classification of ECG beats," *Digital Signal Processing*, vol. 19, no. 2, pp. 320–329, 2009.
- [8] T. Froese, S. Hadjiloucas, R. K. H. Galvão, V. M. Becerra, and C. J. Coelho, "Comparison of extrasystolic ECG signal classifiers using discrete wavelet transforms," *Pattern Recognition Letters*, vol. 27, no. 5, pp. 393–407, 2006.
- [9] S. N. Yu and K. T. Chou, "Integration of independent component analysis and neural networks for ECG beat classification," *Expert Systems with Applications*, vol. 34, no. 4, pp. 2841–2846, 2008.
- [10] J.S. Wang, W.C. Chiang, Y.L. Hsu, and Y.T. C. Yang, "ECG arrhythmia classification using a probabilistic neural network with a feature reduction method," *Neurocomputing*, vol. 116, pp. 38–45, 2013.
- [11] A. Gacek, "Preprocessing and analysis of ECG signals - A self-organizing maps approach," *Expert Systems with Applications*, vol. 38, no. 7, pp. 9008–9013, 2011.

- [12] M. M. Al Rahhal, Y. Bazi, H. Alhichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," *Information Sciences*, vol. 345, pp. 340–354, 2016.
- [13] S. N. Yu and K. T. Chou, "Selection of significant independent components for ECG beat classification," *Expert Systems with Applications*, vol. 36, no. 2 PART 1, pp. 2088–2096, 2009.
- [14] E. Zellmer, F. Shang, and H. Zhang, "Highly Accurate ECG Beat Classification Based on Continuous Wavelet Transformation and Multiple Support Vector Machine Classifiers," in *2009 2nd International Conference on Biomedical Engineering and Informatics*, 2009, pp. 1–5.
- [15] Z. Dokur and T. Ölmez, "ECG beat classification by a novel hybrid neural network," *Computer Methods and Programs in Biomedicine*, vol. 66, no. 2–3, pp. 167–181, 2001.
- [16] U. Rajendra Acharya, P. Subbanna Bhat, S. S. Iyengar, A. Rao, and S. Dua, "Classification of heart rate data using artificial neural network and fuzzy equivalence relation," *Pattern Recognition*, vol. 36, no. 1, pp. 61–68, 2003.
- [17] M. H. Song, J. Lee, S. P. Cho, K. J. Lee, and S. K. Yoo, "Support vector machine-based arrhythmia classification using reduced features," *International Journal of Control, Automation and Systems*, vol. 3, no. 4, pp. 571–579, 2005.
- [18] Y. Kutlu and D. Kuntalp, "A multi-stage automatic arrhythmia recognition and classification system," *Computers in Biology and Medicine*, vol. 41, no. 1, pp. 37–45, 2011.
- [19] R. J. Martis, U. R. Acharya, and H. Adeli, "Current methods in electrocardiogram characterization," *Computers in Biology and Medicine*, vol. 48, no. 1, pp. 133–149, 2014.
- [20] R. J. Martis, U. R. Acharya, K. M. Mandana, A. K. Ray, and C. Chakraborty, "Cardiac decision making using higher order spectra," *Biomedical Signal Processing and Control*, vol. 8, no. 2, pp. 193–203, 2013.
- [21] M. K. Das and S. Ari, "ECG Beats Classification Using Mixture of Features," *International Scholarly Research Notices*, vol. 2014, pp. 1–12, 2014.
- [22] E. D. Übeyli, "Statistics over features of ECG signals," *Expert Systems with Applications*, vol. 36, no. 5, pp. 8758–8767, 2009.
- [23] P. De Chazal and R. B. Reilly, "A patient-adapting heartbeat classifier using ECG morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 12, pp. 2535–2543, 2006.
- [24] N. Acir, "Classification of ECG beats by using a fast least square support vector machines with a dynamic programming feature selection algorithm," *Neural Computing and Applications*, vol. 14, no. 4, pp. 299–309, 2005.

- [25] N. Jannah, S. Hadjiloucas, F. Hwang, and R. K. H. Galvão, "Smart-phone based electrocardiogram wavelet decomposition and neural network classification," *Journal of Physics: Conference Series*, vol. 450, pp. 1–7, 2013.
- [26] P. Bouboulis, S. Theodoridis, C. Mavroforakis, and L. Evaggelatos-Dalla, "Complex Support Vector Machines for Regression and Quaternary Classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 6, pp. 1260 – 1274, 2015.
- [27] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals.," *Circulation*, vol. 101, no. 23, pp. 1–6, 2000.
- [28] N. Jannah and S. Hadjiloucas, "Detection of ECG Arrhythmia Conditions using CSVM and MSVM Classifiers," in *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 2015, pp. 1–2.
- [29] N. Jannah and S. Hadjiloucas, "ECG Arrhythmia Classifiers: Comparison between Multi-class SVM and CSVM Algorithms," in *9th Saudi Students Conference in the UK - 2016 Abstracts of Accepted Papers and Posters*, 2016, p. 481.
- [30] V. Vijaya, K. K. Rao, and V. Rama, "Arrhythmia Detection through ECG Feature Extraction using Wavelet Analysis," *European Journal of Scientific Research*, vol. 66, no. 3, pp. 441–448, 2011.
- [31] D. C. McConahy, "Application of multiobjective optimization to determining an optimal left ventricular assist device (LVAD) pump speed," University of Pittsburgh, 2007.
- [32] A. Timmis and P. Mills, "The Cardiovascular System," *Hutchison's Clinical Methods*, pp. 86–87, 2002.
- [33] M. W. Zimmerman, "Classification of ECG ST events as ischemic or non-ischemic using reconstructed phase spaces," Marquette University, 2004.
- [34] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: Principles and Applications of Bioelectric and Biomagnetic Fields*. New York: Oxford University Press, 1995.
- [35] Texa Health Institute, "Heart Anatomy," *Health Information Center*, 2012. [Online]. Available: <http://www.texasheart.org/HIC/Anatomy/anatomy2.cfm>. [Accessed: 12-Nov-2015].
- [36] A. L. Goldberger, Z. D. Goldberger, and A. Shvilkin, *Goldberger's Clinical Electrocardiography: A Simplified Approach*, 8th ed. Philadelphia: Elsevier Saunders Inc, 2013.
- [37] L. Gavrilovska, S. Krco, V. Milutinovic, I. Stojmenovic, and R. Trobec, *Application and Multidisciplinary Aspects of Wireless Sensor Networks: Concepts, Integration, and Case Studies*. London: Springer-Verlag London Limited, 2011.

- [38] A. Gacek and W. Pedrycz, *ECG Signal Processing, Classification and Interpretation: A Comprehensive Framework of Computational Intelligence*. London: Springer-Verlag London Limited, 2012.
- [39] B. Foster, *Twelve-Lead Electrocardiography*, 2nd ed. London: Springer-Verlag London Limited, 2007.
- [40] Jmarchn, "Limb leads 2 CAT," *Wikimedia*, 2016. [Online]. Available: [https://commons.wikimedia.org/wiki/File:Limb\\_leads\\_2\\_CAT.svg](https://commons.wikimedia.org/wiki/File:Limb_leads_2_CAT.svg). [Accessed: 01-Jan-2016].
- [41] K. M. Lewis and K. A. Handal, *Sensible Analysis of the 12-lead ECG*. Canada: Delmar Thomson Learning, 2000.
- [42] J. R. Hampton, *The ECG made easy*, 8th ed. UK: Churchill Livingstone Elsevier, 2013.
- [43] W. Brady, K. Hundson, R. Naples, A. Sudhir, S. Mitchell, R. Reiser, and J. Ferguson, *ECGs for Prehospital Emergency Care*. UK: Wiley-Blackwell, 2013.
- [44] G. D. Clifford, F. Azuaje, and P. E. McSharry, *Advanced Methods and Tools for ECG Data Analysis*. London Boston: Artech House, INC, 2006.
- [45] J. L. R. Sotelo, "Biosignal analysis for cardiac arrhythmia detection using non-supervised techniques," Universidad Nacional de Colombia, 2010.
- [46] S.T.Sanamdikar, S.T.Hamde, "A Literature Review on Arrhythmia Analysis of ECG Signal," *International Research Journal of Engineering and Technology*, vol. 2, no. 3, pp. 307–312, 2015.
- [47] V. Kumar, A. K. Abbas, N. Fausto, and J. c Aster, *Robbins and Coltron Pathologic Basis of Disease*, 8th ed. United States of America: Philadelphia: Saunders, an imprint of Elsevier Inc, 2010.
- [48] F. J. Chin, "a Real Time Data Mining Technique Applied for Critical ECG Rhythm on Handheld Device," RMIT University, 2012.
- [49] D. R. Karius, "The Physiological Basis of the Abnormal ECG Bundle Branch Blocks," *Kansas City University of Medicine & Biosciences*, 2015. .
- [50] G. Breithardt and O.A. Breithardt, "Left Bundle Branch Block, an Old–New Entity," *Journal of Cardiovascular Translational Research*, vol. 5, no. 2. pp. 107–116, 2012.
- [51] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.
- [52] E. N. Bruce, *Biomedical Signal Processing and Signal Modeling*. New York :Wiley, 2001.
- [53] R. M. Rangayyan, *Biomedical signal analysis : a case-study approach*. United States: IEEE Press, 2002.

- [54] N. V. Thakor and Y.-S. S. Zhu, "Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection," *IEEE Transactions on Biomedical Engineering*, vol. 38, no. 8, pp. 785–794, 1991.
- [55] B. Arvinti, M. Costache, D. Toader, M. Oltean, and A. Isar, "ECG statistical denoising in the wavelet domain," *2010 9th International Symposium on Electronics and Telecommunications, ISETC'10 - Conference Proceedings*, pp. 307–310, 2010.
- [56] A. Velayudhan and P. Soniya, "Study of Different ECG Signal Denoising Techniques," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 3, no. 8, pp. 7733–7739, 2015.
- [57] P. Khandait, N. Bawane, and S. Limaye, "Features Extraction of ECG signal for Detection of Cardiac Arrhythmias," *International Journal of Computer Applications*, , vol. 2, no. 3, pp. 6–10, 2012.
- [58] H. M. Rai and A. Trivedi, "Classification of ECG Waveforms for Abnormalities Detection using DWT and Back Propagation Algorithm," *International Journal of Advanced Research in Computer Engineering and Technology*, vol. 1, no. 4, pp. 517–520, 2012.
- [59] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelet Toolbox™ User's Guide*. The MathWorks, Inc., 2016.
- [60] P. De Chazal, B. G. Celler, and R. B. Reilly, "Using wavelet coefficients for the classification of the ECG Electrocardiogram," in *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2000, vol. 1, pp. 64–67.
- [61] J. Kuzilek, L. Lhotska, and M. Hanuliak, "An automatic method for holter ECG denoising using ICA," *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies - ISABEL '11*, pp. 1–5, 2011.
- [62] M. Popescu, P. Cristea, and A. Bezerianos, "High resolution ECG filtering using adaptive Bayesian wavelet shrinkage," *Computers in Cardiology 1998*, vol. 25, pp. 401–404, 1998.
- [63] Y. Kopsinis and S. McLaughlin, "Development of EMD-based denoising methods inspired by wavelet thresholding," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1351–1362, 2009.
- [64] M. Alfaouri and K. Daqrouq, "ECG signal denoising by wavelet transform thresholding," *American Journal of Applied Sciences*, vol. 5, no. 3, pp. 276–281, 2008.
- [65] P. S. Gokhale, "ECG Signal De-noising using Discrete Wavelet Transform for removal of 50Hz PLI noise," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 5, pp. 81–85, 2012.

- [66] G. Georgieva-tsaneva and K. Tcheshmedjiev, "Denoising of Electrocardiogram Data with Methods of Wavelet Transform," in *International Conference on Computer Systems and Technologies - CompSysTech'13*, 2013, pp. 9–16.
- [67] M. Islam and G. Sajjad, "Performance Comparison of Modified LMS and RLS Algorithms in De-noising of ECG Signals," *International Journal of Engineering and Technology*, vol. 2, no. 3, pp. 466–468, 2012.
- [68] U. Biswas, "Power Line Interference Removal from ECG Signal Using Notch Filter , Adaptive Filter and Wavelet Packet Transform," *American International Journal of Contemporary Scientific Research*, vol. 2, no. 2, pp. 17–23, 2015.
- [69] M. Kaur and B. Singh, "Powerline interference reduction in ECG using combination of MA method and IIR notch," *International Journal of Recent Trends in Engineering*, vol. 2, no. 6, pp. 2–6, 2009.
- [70] C. Chandrakar and M. K. Kowar, "Denoising ECG Signals Using Adaptive Filter Algorithm," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 2, no. 1, pp. 120–123, 2012.
- [71] A. B. Sankar, D. Kumar, and K. Seethalakshmi, "Performance Study of Various Adaptive filter algorithms for Noise Cancellation in Respiratory Signals," *Signal Processing : An International Journal (SPIJ)*, vol. 4, no. 5, pp. 267 – 278, 2010.
- [72] M. Üstündağ, M. Gökbulut, A. Şengür, and F. Ata, "Denoising of weak ECG signals by using wavelet analysis and fuzzy thresholding," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 1, no. 4, pp. 135–140, 2012.
- [73] M. A. Kabir and C. Shahnaz, "Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains," *Biomedical Signal Processing and Control*, vol. 7, no. 5, pp. 481–489, 2012.
- [74] M. AlMahamdy and H. B. Riley, "Performance study of different denoising methods for ECG signals," *Procedia Computer Science*, vol. 37, pp. 325–332, 2014.
- [75] S. Gang, R. Liang, and G. Yuhan, "Research on the Method of ECG Denoising Based on Improved Wavelet Packet," in *2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, 2015, pp. 1425–1428.
- [76] P. Phukpattaranont, "QRS detection algorithm based on the quadratic filter," *Expert Systems with Applications*, vol. 42, no. 11, pp. 4867–4877, 2015.
- [77] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm.," *IEEE transactions on bio-medical engineering*, vol. 32, no. 3, pp. 230–236, 1985.
- [78] Hamilton P. and Tompkins W., "Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database," *IEEE Transactions in Biomedical Engineering*, vol. BME-33, no. 12, pp. 1157–1165, 1986.

- [79] B. T. Krishna, "QRS detection using digital differentiators," in *2015 International Conference on Electrical Engineering and Informatics (ICEEI)*, 2015, pp. 356–359.
- [80] D. Benitez, P. A. Gaydecki, A. Zaidi, and A. P. Fitzpatrick, "The use of the Hilbert transform in ECG signal analysis," *Computers in Biology and Medicine*, vol. 31, no. 5, pp. 399–406, 2001.
- [81] I. A. Tarmizi, S. S. N. A. S. Hassan and W. P. W. Ibrahim, "A journal of real peak recognition of electrocardiogram (ECG) signals using neural network," in *2012 2nd International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, 2012, pp. 504–510.
- [82] Y. Xia, J. Han, and K. Wang, "Quick detection of QRS complexes and R-waves using a wavelet transform and K-means clustering," *Bio-Medical Materials and Engineering*, vol. 26, pp. S1059–S1065, 2015.
- [83] S. Choudhary and S. T. Hamde, "A simple and robust algorithm for the detection of QRS complexes," in *2015 International Conference on Industrial Instrumentation and Control (ICIC)*, 2015, pp. 870–872.
- [84] L. D. Sharma and R. K. Sunkaria, "A robust QRS detection using novel pre-processing techniques and kurtosis based enhanced efficiency," *Measurement*, vol. 87, pp. 194–204, 2016.
- [85] J. Arteaga-falconi, H. Al Osman, and A. El Saddik, "R-Peak Detection Algorithm Based on Differentiation," in " *2015 IEEE 9th International Symposium on Intelligent Signal Processing (WISP) Proceedings*, 2015, pp. 1–4.
- [86] D. Lai, F. Zhang, and C. Wang, "A Real-time QRS Complex Detection Algorithm Based on Differential Threshold Method," in *2015 IEEE International Conference on Digital Signal Processing (DSP)*, 2015, pp. 129–133.
- [87] R. J. Martis, U. R. Acharya, and L. C. Min, "ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform," *Biomedical Signal Processing and Control*, vol. 8, no. 5, pp. 437–448, 2013.
- [88] F. A. Elhaj, N. Salim, A. R. Harris, T. T. Swee, and T. Ahmed, "Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals," *Computer Methods and Programs in Biomedicine*, vol. 127, pp. 52–63, 2016.
- [89] S. Ayub and J. Saini, "ECG classification and abnormality detection using cascade forward neural network," *International Journal of Engineering, Science and Technology*, vol. 3, no. 3, pp. 41–46, 2011.
- [90] A. Rabee and I. Barhumi, "ECG signal classification using support vector machine based on wavelet multiresolution analysis," in *The 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, 2012, pp. 1319–1323.

- [91] I. Jekova, G. Bortolan, and I. Christov, "Assessment and comparison of different methods for heartbeat classification," *Medical Engineering and Physics*, vol. 30, no. 2, pp. 248–257, 2008.
- [92] Y. H. Hu, S. Palreddy, and W. J. Tompkins, "A patient-adaptable ECG beat classifier using a mixture of experts approach.," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 9, pp. 891–900, 1997.
- [93] K. I. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 2, pp. 179–185, 1999.
- [94] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "Multirate processing of the ECG using filter banks," *Computers in Cardiology*, pp. 245–248, 1996.
- [95] S. Osowski and T. H. Linh, "ECG beat recognition using fuzzy hybrid neural network," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 11, pp. 1265–1271, 2001.
- [96] A. K. M. F. Haque, H. Ali, M. A. Kiber, and T. Hasan, "Automatic Feature Extraction of ECG Signal Using Fast Fourier Transform," 2009, [Online]. Available: [https://www.researchgate.net/profile/Dr\\_AKM\\_Fazlul\\_Haque/publication/295813369\\_Automatic\\_Feature\\_Extraction\\_of\\_ECG\\_Signal\\_Using\\_Fast\\_Fourier\\_Transform/links/56cda7df08ae85c8233e5fe6.pdf](https://www.researchgate.net/profile/Dr_AKM_Fazlul_Haque/publication/295813369_Automatic_Feature_Extraction_of_ECG_Signal_Using_Fast_Fourier_Transform/links/56cda7df08ae85c8233e5fe6.pdf). [Accessed: 12-Dec-2015].
- [97] H. Gothwal, "Cardiac arrhythmias detection in an ECG beat signal using fast fourier transform and artificial neural network," *Journal of Biomedical Science and Engineering*, vol. 04, no. 04, pp. 289–296, 2011.
- [98] T. H. Linh, S. Osowski, and M. Stodolski, "On-line heart beat recognition using hermite polynomials and neuro-fuzzy network," *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 4, pp. 1224–1231, 2003.
- [99] S. N. Yu and K. T. Chou, "A switchable scheme for ECG beat classification based on independent component analysis," *Expert Systems with Applications*, vol. 33, no. 4, pp. 824–829, 2007.
- [100] M. I. Owis, A. H. Abou-Zied, A. B. M. Youssef, and Y. M. Kadah, "Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 7, pp. 733–736, 2002.
- [101] E. D. Übeyli, "ECG beats classification using multiclass support vector machines with error correcting output codes," *Digital Signal Processing*, vol. 17, no. 3, pp. 675–684, 2007.
- [102] Y. Özbay, R. Ceylan, and B. Karlik, "Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier," *Expert Systems with Applications*, vol. 38, no. 1, pp. 1004–1010, 2011.

- [103] F. M. Vaneghi, M. Oladazimi, F. Shiman, A. Kordi, M. J. Safari, and F. Ibrahim, "A Comparative Approach to ECG Feature Extraction Methods," *2012 Third International Conference on Intelligent Systems Modelling and Simulation*, pp. 252–256, 2012.
- [104] B. M. Asl, S. K. Setarehdan, and M. Mohebbi, "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal," *Artificial Intelligence in Medicine*, vol. 44, no. 1, pp. 51–64, 2008.
- [105] J. S. M. Murugappan, K. Wan, and S. Yaacob, "Electrocardiogram-based emotion recognition system using empirical mode decomposition and discrete Fourier transform," *Expert Systems*, vol. 31, no. 2, pp. 110–120, 2013.
- [106] Y.C. Yeh, W.J. Wang, and C. W. Chiou, "Cardiac arrhythmia diagnosis method using linear discriminant analysis on ECG signals," *Measurement*, vol. 42, no. 5, pp. 778–789, 2009.
- [107] T. H. Yeap, F. Johnson, and M. Rachniowski, "ECG Beat Classification By A Neural Network," in *[1990] Proceedings of the Twelfth Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1990, pp. 1457–1458.
- [108] S. Palreddy and W. J. Tompkins, "Customization of ECG beat classifiers developed using SOM and LVQ," in *Proceedings of 17th International Conference of the Engineering in Medicine and Biology Society*, 1995, vol. 1, pp. 813–814.
- [109] Y. Özbay, R. Ceylan, and B. Karlik, "A fuzzy clustering neural network architecture for classification of ECG arrhythmias," *Computers in Biology and Medicine*, vol. 36, no. 4, pp. 376–388, 2006.
- [110] R. Sathya and A. Abraham, "Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification," *International Journal of advanced Research in Artificial Intelligence*, vol. 2, no. 2, pp. 34–38, 2013.
- [111] E. Alegre, V. González-Castro, S. Suárez, and M. Castejón, "Comparison of supervised and unsupervised methods to classify boar acrosomes using texture descriptors," in *2009 International Symposium ELMAR*, 2009, pp. 65–70.
- [112] M. H. Beale, M. T. Hagan, and B. H. Demuth, *Neural Network Toolbox User 's Guide How to*. The MathWorks, Inc., 2015.
- [113] M. Korürek and B. Doğan, "ECG beat classification using particle swarm optimization and radial basis function neural network," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7563–7569, 2010.
- [114] T. Stamkopoulos, K. Diamantaras, N. Maglaveras, and M. Srintzis, "ECG analysis using nonlinear PCA neural networks for ischemia detection," *IEEE Transactions on Signal Processing*, vol. 46, no. 11, pp. 3058–3067, 1998.
- [115] D. Gao, M. Madden, D. Chambers, and G. Lyons, "Bayesian ANN classifier for ECG arrhythmia diagnostic system: A comparison study," *Proceedings of the International Joint Conference on Neural Networks*, vol. 4, pp. 2383–2388, 2005.

- [116] O. Sayadi, M. B. Shamsollahi, and G. D. Clifford, "Robust detection of premature ventricular contractions using a wave-based Bayesian framework.," *IEEE transactions on bio-medical engineering*, vol. 57, no. 2, pp. 353–362, 2010.
- [117] R. Kazemi, A. Farsi, M. H. Ghaed, and M. Karimi-Ghartemani, "Detection and extraction of periodic noises in audio and biomedical signals using Kalman filter," *Signal Processing*, vol. 88, no. 8, pp. 2114–2121, 2008.
- [118] Y. P. Meau, F. Ibrahim, S. A. L. Narainasamy, and R. Omar, "Intelligent classification of electrocardiogram (ECG) signal using extended Kalman Filter (EKF) based neuro fuzzy system," *Computer Methods and Programs in Biomedicine*, vol. 82, no. 2, pp. 157–168, 2006.
- [119] U. Aydin and Y. S. Dogrusoz, "A Kalman filter-based approach to reduce the effects of geometric errors and the measurement noise in the inverse ECG problem," *Medical and Biological Engineering and Computing*, vol. 49, no. 9, pp. 1003–1013, 2011.
- [120] M. Engin, "ECG beat classification using neuro-fuzzy network," *Pattern Recognition Letters*, vol. 25, no. 15, pp. 1715–1722, 2004.
- [121] R. Ceylan, Y. Özbay, and B. Karlik, "A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network," *Expert Systems with Applications*, vol. 36, no. 3 PART 2, pp. 6721–6726, 2009.
- [122] A. Sengur and I. Turkoglu, "A hybrid method based on artificial immune system and fuzzy k-NN algorithm for diagnosis of heart valve diseases," *Expert Systems with Applications*, vol. 35, no. 3, pp. 1011–1020, 2008.
- [123] E. Ramirez, O. Castillo, and J. Soria, "Hybrid system for cardiac arrhythmia classification with fuzzy k-nearest neighbors and Multi Layer Perceptrons combined by a fuzzy inference system," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*, 2010, pp. 1–6.
- [124] F. Bagheri, N. Ghafarnia, and F. Bahrami, "Electrocardiogram (ECG) Signal Modeling and Noise Reduction Using Hopfield Neural Networks," *Engineering, Technology and Applied Science Research*, vol. 3, no. 1, pp. 345–348, 2013.
- [125] E. Derya Übeyli, "Recurrent neural networks employing Lyapunov exponents for analysis of ECG signals," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1192–1199, 2010.
- [126] J. C. Gallagher, S. K. Boddhu, and S. Vignraham, "A reconfigurable continuous time recurrent neural network for evolvable hardware applications," in *Proceedings - NASA/DoD Conference on Evolvable Hardware, (EH'05)*, 2005, pp. 247–250.
- [127] J. C. Gallagher and J. M. Fiore, "Continuous Time Recurrent Neural Networks: a Paradigm for Evolvable Analog Controller Circuits," in *The Proceedings of the National Aerospace and Electronics Conference 2000 (NAECON) 10 -12 Oct 2000*, 2000, pp. 299–304.

- [128] A. Seyab, R. K. Shakir, and Y. Cao, "Nonlinear system identification for predictive control using continuous time recurrent neural networks and automatic differentiation," *Journal of Process Control*, vol. 18, no. 6, pp. 568–581, 2008.
- [129] I. Güler and E. D. Übeyli, "ECG beat classifier designed by combined neural network model," *Pattern Recognition*, vol. 38, no. 2, pp. 199–208, 2005.
- [130] G. K. Prasad and J. S. Sahambi, "Classification of ECG arrhythmias using multi-resolution analysis and neural networks," in *TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region*, 2003, vol. 1, pp. 227–231.
- [131] E. D. Übeyli and I. Güler, "Detection of electrocardiographic changes in partial epileptic patients using Lyapunov exponents with multilayer perceptron neural networks," *Engineering Applications of Artificial Intelligence*, vol. 17, no. 6, pp. 567–576, 2004.
- [132] I. Güler and E. D. Übeyli, "An expert system for detection of electrocardiographic changes in patients with partial epilepsy using wavelet-based neural networks," *Expert Systems*, vol. 22, no. 2, pp. 62–71, 2005.
- [133] S. S. Mehta and N. S. Lingayat, "Combined entropy based method for detection of QRS complexes in 12-lead electrocardiogram using SVM," *Computers in Biology and Medicine*, vol. 38, no. 1, pp. 138–145, 2008.
- [134] S. S. Mehta and N. S. Lingayat, "Development of SVM based classification techniques for the delineation of wave components in 12-lead electrocardiogram," *Biomedical Signal Processing and Control*, vol. 3, no. 4, pp. 341–349, 2008.
- [135] D. Sambhu and A. C. Umesh, "Automatic Classification of ECG Signals with Features Extracted Using Wavelet Transform and Support Vector Machines," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 2, no. 1, pp. 235–241, 2013.
- [136] H. Zhang and L.-Q. Zhang, "ECG analysis based on PCA and Support Vector Machines," in *2005 International Conference on Neural Networks and Brain*, 2005, vol. 2, pp. 743–747.
- [137] R. J. Martis, U. R. Acharya, K. M. Mandana, a. K. Ray, and C. Chakraborty, "Application of principal component analysis to ECG signals for automated diagnosis of cardiac health," *Expert Systems with Applications*, vol. 39, no. 14, pp. 11792–11800, 2012.
- [138] S. Karpagachelvi, M. Arthanari, and M. Sivakumar, "Classification of ECG Signals Using Extreme Learning Machine," *Computer and Information Science*, vol. 4, no. 1, pp. 42–52, 2011.
- [139] M. Lagerholm and G. Peterson, "Clustering ECG complexes using hermite functions and self-organizing maps," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 7, pp. 838–848, 2000.

- [140] M. Gaetz, G. L. Iverson, E. J. Rzempoluck, R. Remick, P. McLean, and W. Linden, "Self-organizing neural network analyses of cardiac data in depression," *Neuropsychobiology*, vol. 49, no. 1, pp. 30–37, 2004.
- [141] M. L. Talbi and A. Charef, "PVC discrimination using the QRS power spectrum and self-organizing maps," *Computer Methods and Programs in Biomedicine*, vol. 94, no. 3, pp. 223–231, 2009.
- [142] H. Nishizawa, T. Obi, M. Yamaguchi, and N. Ohyama, "Hierarchical Clustering Method for Extraction of Knowledge from a Large Amount of Data," *Optical Review*, vol. 6, no. 4, pp. 302–307, 1999.
- [143] S. Boudaoud, H. Rix, J. J. Blanc, J. C. Cornily, and O. Meste, "Integrated shape averaging of the P-wave applied to AF risk detection," in *Computers in Cardiology, 2003*, 2003, pp. 125–128.
- [144] The MathWorks, "MATLAB: The Language of Technical Computing, Getting Started with MATLAB," 2001. [Online]. Available: <https://web.stanford.edu/class/ee262/software/getstart.pdf>. [Accessed: 12-Dec-2015].
- [145] The MathWorks., *Signal Processing Toolbox<sup>TM</sup> User's Guide*. The MathWorks, Inc., 2016.
- [146] C. Chang and C. Lin, "LIBSVM: A Library for Support Vector Machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, pp. 1–39, 2011.
- [147] C. W. Hsu, C. C. Chang, and C. J. Lin, "A Practical Guide to Support Vector Classification," 2010.
- [148] I. Silva and G. B. Moody, "An Open-source Toolbox for Analysing and Processing PhysioNet Databases in MATLAB and Octave," *Journal of Open Research Software*, vol. 2, no. 1, pp. 1–7, 2014.
- [149] Y. C. Yeh and W. J. Wang, "QRS complexes detection for ECG signal: The Difference Operation Method," *Computer Methods and Programs in Biomedicine*, vol. 91, no. 3, pp. 245–254, 2008.
- [150] Q. Xue, Y. H. Hu, and W. J. Tompkins, "Neural-Network-Based Adaptive Matched Filtering for QRS Detection," *IEEE Transactions on Biomedical Engineering*, vol. 39, no. 4, pp. 317–329, 1992.
- [151] S. Pal and M. Mitra, "Detection of ECG characteristic points using Multiresolution Wavelet Analysis based Selective Coefficient Method," *Measurement: Journal of the International Measurement Confederation*, vol. 43, no. 2, pp. 255–261, 2010.
- [152] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989.

- [153] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "ECG beat detection using filter banks," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 2, pp. 192–202, 1999.
- [154] N. Debbabi and S. El Asmi, "Algebraic approach for R-peak detection in the ElectroCardioGram (ECG) signal," *ICSCS 2012 - 2012 1st International Conference on Systems and Computer Science*, pp. 1–5, 2012.
- [155] C. Papaloukas, D. I. Fotiadis, A. Likas, and L. K. Michalis, "Automated methods for ischemia detection in long-duration ECGs," Greece, 2002.
- [156] K. Qureshi and V. P. Patel, "Efficient Data Compression of Ecg Signal Using Discrete Wavelet Transform," *International Journal of Research in Engineering and Technology (IJRET)*, vol. 2, no. 4, pp. 696–699, 2013.
- [157] A. F. Ahmed, "A Comparative Study of ECG Signal Compression Techniques Using Various Transforms," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 5, no. 6, pp. 60–64, 2015.
- [158] K. R. Rao and P. Yip, *Discrete cosine transform: algorithms, advantages, applications*. San Diego: Academic Press Professional, Inc., 1990.
- [159] V. K. Ingle and J. G. Proakis, *Digital Signal Processing Using MATLAB*, 3rd ed. United States of America: Cengage Learning, 2012.
- [160] S. C. Chan and K. L. Ho, "Direct methods for computing discrete sinusoidal transforms," *IEE Proceedings F-Radar and Signal Processing*, vol. 137, no. 6, pp. 433–442, 1990.
- [161] A. Ashok, N. Agarwal, M. Kumar, A. K. Jaiswal, and K. S. Sachan, "Performance Analysis of Various Types of ECG Compression Techniques In Terms of CR," *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, vol. 1, no. 1, pp. 1–5, 2012.
- [162] V. Britanak, P. C. Yip, and K. Rao, *Discrete Cosine and Sine Transforms: General Properties, Fast Algorithms and Integer Approximations*. Amsterdam: Academic Press Inc., Elsevier Science, 2007.
- [163] R. K. R. Yarlagadda, *Analog and Digital Signals and Systems*. U.S.A: Springer US, 2010.
- [164] R. Kumar, A. Kumar, and G. K. Singh, "Hybrid method based on singular value decomposition and embedded zero tree wavelet technique for ECG signal compression," *Computer Methods and Programs in Biomedicine*, vol. 129, pp. 135–148, 2015.
- [165] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley-Interscience, 2001.
- [166] V. Gupta, R. Singh, G. Singh, R. Singh, and H. Singh, "An Introduction to Principal Component Analysis and Its Importance in Biomedical Signal Processing," 2011

- International Conference on Life Science and Technology IPCBEE*, vol. 3, pp. 29–33, 2011.
- [167] F. Vargas, D. Lettnin, M. C. F. De Castro, and M. Macarthy, “Electrocardiogram pattern recognition by means of MLP network and PCA: A case study on equal amount of input signal types,” in *VII Brazilian Symposium on Neural Networks, 2002. SBRN 2002. Proceedings.*, 2002, pp. 200–205.
- [168] D. Patra, M. K. Das, and S. Pradhan, “Integration of Fcm, Pca and neural networks for classification of Ecg arrhythmias,” *IAENG International Journal of Computer Science*, vol. 36, no. 3, pp. 1–5, 2010.
- [169] L. Kanaan, D. Merheb, M. Kallas, C. Francis, H. Amoud, and P. Honeine, “PCA and KPCA of ECG signals with binary SVM classification,” in *2011 IEEE Workshop on Signal Processing Systems (SiPS)*, 2011, pp. 344–348.
- [170] R. Jain, “12 Lead ECG Data Compression Using Principal Component Analysis,” *International Journal of Emerging Trends and Technology in Computer Science (IJETTCS)*, vol. 4, no. 1, pp. 34–39, 2015.
- [171] J. Shlens, “A Tutorial on Principal Component Analysis,” *arXiv preprint arXiv:1404.1100*, 2014.
- [172] R. Silipo and C. Marchesi, “Artificial neural networks for automatic ECG analysis,” *IEEE Transactions on Signal Processing*, vol. 46, no. 5, pp. 1417–1425, 1998.
- [173] A. K. Jain, J. Mao, and K. M. Mohiuddin, “Artificial neural networks: a tutorial,” *Computer*, vol. 29, no. 3, pp. 31–44, 1996.
- [174] K. Polat and S. Güneş, “Detection of ECG Arrhythmia using a differential expert system approach based on principal component analysis and least square support vector machine,” *Applied Mathematics and Computation*, vol. 186, no. 1, pp. 898–906, 2007.
- [175] R. Ceylan and Y. Özbay, “Comparison of FCM, PCA and WT techniques for classification ECG arrhythmias using artificial neural network,” *Expert Systems with Applications*, vol. 33, no. 2, pp. 286–295, 2007.
- [176] D. L. Olson and D. Delen, *Advanced Data Mining Techniques*. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg, 2008.
- [177] V. Vapnik, *Estimation of Dependences Based on Empirical Data*, 2nd ed. US: Springer, 2006.
- [178] E. Osuna, R. Freund, and F. Girosi, “An improved training algorithm for support vector machines,” in *Neural Networks for Signal Processing [1997] VII. Proceedings of the 1997 IEEE Workshop*, 1997, pp. 276–285.
- [179] S. S. Keerthi, S. K. Shevade, C. Bhattacharyya, and K. R. K. Murthy, “Improvements to Platt’s SMO Algorithm for SVM Classifier Design,” *Neural Computation*, vol. 13, no. 3, pp. 637–649, 2001.

- [180] T. Fletcher, "Support Vector Machines Explained," 2009. [Online]. Available: <http://sutikno.blog.undip.ac.id/files/2011/11/SVM-Explained.pdf>. [Accessed: 06-Jun-2014].
- [181] G. G. Sanchez, "Examination of the applicability of Support Vector Machines in the context of ischaemia detection," *UPCs new academic and scientific communications*, 2011.
- [182] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. New York: Cambridge University Press, 2004.
- [183] O. Carugo and F. Eisenhaber, *Data Mining Techniques for the Life Sciences*. New York: Humana Press, 2010.
- [184] S. R. Gunn, "Support Vector Machines for Classification and Regression," *Image Speech and Intelligent Systems (ISIS), Technical Report*, vol. 14, 1998.
- [185] S. V Albrecht, "SVM Toolbox, Theory, Documentation, Experiments," *Technische Universitat Darmstadt*, 2010 .
- [186] S. Manandhar, J. Austin, U. Desai, Y. Oyanagi, and A. Talukder, *Applied Computing: Second Asian Applied Computing Conference, AACC 2004, Kathmandu, Nepal, October 29-31, 2004. Proceedings*. New York: Springer Berlin Heidelberg, 2004.
- [187] S. Abe and S. Singh, *Support Vector Machines for Pattern Classification*, 2nd ed. London: Springer-Verlag, 2010.
- [188] J. C. Platt, "Sequential minimal optimization: A fast algorithm for training support vector machines," *Microsoft Research*, 1998. [Online]. Available: <https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr-98-14.pdf>. [Accessed: 01-Jan-2016].
- [189] P. Bouboulis and S. Theodoridis, "Extension of Wirtinger 's Calculus in Reproducing Kernel Hilbert Spaces and the Complex Kernel LMS Extension of Wirtinger's Calculus in Reproducing Kernel Hilbert Spaces and the Complex Kernel LMS," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 964–978, 2011.
- [190] J. A. Siegel, *Forensic Chemistry: Fundamentals and Applications*. USA: John Wiley & Sons, Ltd., 2016.
- [191] J. D. Bronzino, *The Biomedical Engineering HandBook (Vol. 2)*, 2nd ed. New York: CRC Press, 2000.
- [192] C.-H. Lin, "Frequency-domain features for ECG beat discrimination using grey relational analysis-based classifier," *Computers and Mathematics with Applications*, vol. 55, no. 4, pp. 680–690, 2008.
- [193] S. Deshpande and S. O. Rajankar, "ECG Data Compression Using Principal Component Analysis," *International Journal of Electrical, Electronics and Computer Systems, (IJECS)*, vol. 1, no. 2, pp. 13–16, 2013.

- [194] V. Aggarwal and M. Patterh, "ECG Compression using Wavelet Packet, Cosine Packet and Wave Atom Transforms.," *International Journal of Electronic Engineering Research*, vol. 1, no. 3, pp. 259–268, 2009.
- [195] H. Saberhari and M. Shamsi, "Comparison of different algorithms for ECG signal compression based on transfer coding," *ISIEA 2012 - 2012 IEEE Symposium on Industrial Electronics and Applications*, pp. 341–345, 2012.
- [196] Y. Zigel, A. Cohen, and A. Katz, "ECG signal compression using analysis by synthesis coding," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 10, pp. 1308–16, 2000.
- [197] M. K. Chhipa, "Performance Analysis of Various Transforms Based Methods for ECG Data," *International Journal of Scientific and Research Publications*, vol. 3, no. 5, pp. 1–6, 2013.
- [198] A. K. Agrawal, A. Mishra, and R. Kumar, "ECG Signal Compression using efficient transformations," *International Journal of Innovations in Engineering and Technology (IJJET)*, vol. 4, no. 1, pp. 13–19, 2014.
- [199] N. A. Bhaskar, "Performance analysis of Support Vector Machine and Neural Networks in detection of Myocardial Infarction," *Procedia Computer Science*, vol. 46, pp. 20–30, 2015.
- [200] V. Rathikarani and P. Dhanalakshmi, "Automatic Classification of ECG signal for Identifying Arrhythmia," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 9, pp. 205–211, 2013.
- [201] T. P. Exarchos, C. Papaloukas, D. I. Fotiadis, and L. K. Michalis, "An association rule mining-based methodology for automated detection of ischemic ECG beats," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 8, pp. 1531–1540, 2006.
- [202] A. Vishwa, M. K. Lal, S. Dixit, and P. Vardwaj, "Clasification Of Arrhythmic ECG Data Using Machine Learning Techniques," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 1, no. 4, pp. 68–71, 2011.
- [203] A. De Gaetano, S. Panunzi, F. Rinaldi, A. Risi, and M. Sciandrone, "A patient adaptable ECG beat classifier based on neural networks," *Applied Mathematics and Computation*, vol. 213, no. 1, pp. 243–249, 2009.
- [204] S. Kiranyaz, T. Ince, J. Pulkkinen, and M. Gabbouj, "Personalized long-term ECG classification: A systematic approach," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3220–3226, 2011.
- [205] M. Korürek and A. Nizam, "Clustering MIT-BIH arrhythmias with Ant Colony Optimization using time domain and PCA compressed wavelet coefficients," *Digital Signal Processing: A Review Journal*, vol. 20, no. 4, pp. 1050–1060, 2010.

- [206] S. Hadjiloucas, N. Jannah, F. Hwang, and R. K. H. Galvão, "On the application of optimal wavelet filter banks for ECG signal classification," *Journal of Physics: Conference Series*, vol. 490, pp. 1–5, 2014.
- [207] S. Banerjee, R. Gupta, and M. Mitra, "Delineation of ECG characteristic features using multiresolution wavelet analysis method," *Measurement*, vol. 45, no. 3, pp. 474–487, 2012.
- [208] S. Pal and M. Mitra, "Detection of ECG characteristic points using Multiresolution Wavelet Analysis based Selective Coefficient Method," *Measurement*, vol. 43, no. 2, pp. 255–261, 2010.
- [209] H. H. de Carvalho, R. L. Moreno, T. C. Pimenta, P. C. Crepaldi, and E. Cintra, "A heart disease recognition embedded system with fuzzy cluster algorithm," *Computer Methods and Programs in Biomedicine*, vol. 110, no. 3, pp. 447–454, 2013.
- [210] N. Kohli and N. Verma, "Arrhythmia classification using SVM with selected features," *International Journal of Engineering, Science and Technology*, vol. 3, no. 8, pp. 122–131, 2011.
- [211] N. Acir, "A support vector machine classifier algorithm based on a perturbation method and its application to ECG beat recognition systems," *Expert Systems with Applications*, vol. 31, no. 1, pp. 150–158, 2006.
- [212] K. T. Chou and S. N. Yu, "Categorizing Heartbeats by Independent Component Analysis and Support Vector Machines," *2008 Eighth International Conference on Intelligent Systems Design and Applications*, vol. 1, pp. 599–602, 2008.
- [213] J. A. Nasiri, M. Naghibzadeh, H. S. Yazdi, and B. Naghibzadeh, "ECG arrhythmia classification with support vector machines and genetic algorithm," in *Third UKSim European Symposium on Computer Modelling and Simulation*, 2009, pp. 187–192.
- [214] S. Dutta, A. Chatterjee, and S. Munshi, "Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification," *Medical Engineering and Physics*, vol. 32, no. 10, pp. 1161–1169, 2010.
- [215] D. Thanapatay, C. Suwansaroj, and C. Thanawattano, "ECG beat classification method for ECG printout with Principle Components Analysis and Support Vector Machines," in *2010 International Conference on Electronics and Information Engineering (ICEIE)*, 2010, vol. 1, pp. V1–72–V1–75.
- [216] C. P. Shen, W. C. Kao, Y. Y. Yang, M. C. Hsu, Y. T. Wu, and F. Lai, "Detection of cardiac arrhythmia in electrocardiograms using adaptive feature extraction and modified support vector machines," *Expert Systems with Applications*, vol. 39, no. 9, pp. 7845–7852, 2012.
- [217] R. J. Martis, U. R. Acharya, and L. C. Min, "ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform," *Biomedical Signal Processing and Control*, vol. 8, no. 5, pp. 437–448, 2013.

- [218] R. J. Martis, U. R. Acharya, C. M. Lim, and J. S. Suri, "Characterization of ECG beats from cardiac arrhythmia using discrete cosine transform in PCA framework," *Knowledge-Based Systems*, vol. 45, pp. 76–82, 2013.
- [219] A. Jovic and N. Bogunovic, "Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features," *Artificial Intelligence in Medicine*, vol. 51, no. 3, pp. 175–186, 2011.
- [220] R. Ahmed and S. Arafat, "Cardiac arrhythmia classification using hierarchical classification model," in *6th International Conference on Computer Science and Information Technology (CSIT), 2014*, 2014, pp. 203–207.
- [221] E. Alickovic and A. Subasi, "Effect of Multiscale PCA De-noising in ECG Beat Classification for Diagnosis of Cardiovascular Diseases," *Circuits, Systems, and Signal Processing*, vol. 34, no. 2, pp. 513–533, 2015.
- [222] D. Rezgüi and Z. Lachiri, "ECG biometric recognition using SVM-based approach," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 11, no. S1, pp. S94–S100, 2016.
- [223] M. Hendel and a. Benyettou, "Heartbeats arrhythmia classification using probabilistic multi-class support vector machines: A comparative study," *International Journal of Applied Engineering Research*, vol. 11, no. 2, pp. 1293–1297, 2016.
- [224] I. Saini, D. Singh, and A. Khosla, "Electrocardiogram beat classification using empirical mode decomposition and multiclass directed acyclic graph support vector machine," *Computers and Electrical Engineering*, vol. 40, no. 5, pp. 1774–1787, 2014.
- [225] H. Huang, J. Liu, Q. Zhu, R. Wang, and G. Hu, "Detection of inter-patient left and right bundle branch block heartbeats in ECG using ensemble classifiers.," *Biomedical engineering online*, vol. 13, no. 72, pp. 1–22, 2014.
- [226] İ. Güler and E. D. Übeyli, "Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction," *Expert Systems with Applications*, vol. 27, no. 3, pp. 323–330, 2004.
- [227] K. Polat, S. Şahan, and S. Güneş, "A new method to medical diagnosis: Artificial immune recognition system (AIRS) with fuzzy weighted pre-processing and application to ECG arrhythmia," *Expert Systems with Applications*, vol. 31, no. 2, pp. 264–269, 2006.
- [228] E. J. D. S. Luz, D. Menotti, and W. R. Schwartz, "Evaluating the use of ECG signal in low frequencies as a biometry," *Expert Systems with Applications*, vol. 41, no. 5, pp. 2309–2315, 2014.
- [229] E. J. D. S. Luz, T. M. Nunes, V. H. C. De Albuquerque, J. P. Papa, and D. Menotti, "ECG arrhythmia classification based on optimum-path forest," *Expert Systems with Applications*, vol. 40, no. 9, pp. 3561–3573, 2013.

- [230] M. Moavenian and H. Khorrami, "A qualitative comparison of Artificial Neural Networks and Support Vector Machines in ECG arrhythmias classification," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3088–3093, 2010.
- [231] E. J. Bayro-Corrochano and N. Arana-Daniel, "Clifford support vector machines for classification, regression, and recurrence," *IEEE Transactions on Neural Networks*, vol. 21, no. 11, pp. 1731–1746, 2010.
- [232] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG Analysis : A New Approach in Human Identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808–812, 2001.
- [233] T. M. Nazmy, H. El-Messiry, and B. Al-Bokhity, "Adaptive Neuro-Fuzzy Inference System for Classification OF ECG Signals," *The 7th International Conference on Informatics and Systems (INFOS)*, pp. 1–6, 2010.
- [234] U. Chandrasekhar, L. D. Dwivedi, and C. L. Chowdhary, "Classification of ECG beats using features from two-stage two-band wavelet decomposition," *Journal of Theoretical and Applied Information Technology*, vol. 49, no. 3, pp. 922–928, 2013.
- [235] M. Korürek and A. Nizam, "A new arrhythmia clustering technique based on Ant Colony Optimization," *Journal of Biomedical Informatics*, vol. 41, no. 6, pp. 874–881, 2008.
- [236] A. K. Mishra and S. Raghav, "Local fractal dimension based ECG arrhythmia classification," *Biomedical Signal Processing and Control*, vol. 5, no. 2, pp. 114–123, 2010.
- [237] A. Daamouche, L. Hamami, N. Alajlan, and F. Melgani, "A wavelet optimization approach for ECG signal classification," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 342–349, 2012.
- [238] B. Doğan and M. Korürek, "A new ECG beat clustering method based on kernelized fuzzy c-means and hybrid ant colony optimization for continuous domains," *Applied Soft Computing Journal*, vol. 12, no. 11, pp. 3442–3451, 2012.
- [239] E. Alickovic and A. Subasi, "Medical Decision Support System for Diagnosis of Heart Arrhythmia using DWT and Random Forests Classifier," *Journal of Medical Systems*, vol. 40, no. 4, p. 108, 2016.
- [240] A. A. Tahat, "Mobile personal electrocardiogram monitoring system and transmission using MMS.," in *7th International Caribbean Conference on Devices, Circuits and Systems*, 2008, pp. 1–5.
- [241] C. Wen, M. F. Yeh, K. C. Chang, and R. G. Lee, "Real-time ECG telemonitoring system design with mobile phone platform," *Measurement: Journal of the International Measurement Confederation*, vol. 41, no. 4, pp. 463–470, 2008.
- [242] R. Merzougui and M. Feham, "Algorithm of remote monitoring ECG using mobile phone: Conception and implementation.," in *3rd International Conference on*

- Broadband Communications, Informatics and Biomedical Applications*, 2008, pp. 254–261.
- [243] Q. Fang, F. Sufi, and I. Cosic, “A Mobile Device Based ECG Analysis System,” in *Data Mining in Medical and Biological Research*, Europe: InTech, 2008, pp. 209–226.
- [244] V. Monasterio, P. Laguna, and J. P. Martínez, “Multilead analysis of T-wave alternans in the ecg using principal component analysis,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 7, pp. 1880–1890, 2009.
- [245] L. N. Sharma, S. Danadapat, and A. Mahanta, “Multiscale PCA based Quality Controlled Denoising of Multichannel ECG Signals,” *International Journal of Information and Electronics Engineering*, vol. 2, no. 2, pp. 107–111, 2012.
- [246] R. Almeida, J. P. Martínez, A. P. Rocha, and P. Laguna, “Multilead ECG Delineation Using Spatially Projected Leads From Wavelet Transform Loops,” *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 8, pp. 1996–2005, 2009.
- [247] M. R. Homaeinezhad, A. Ghaffari, and R. Rahmani, “Review: Multi-lead discrete wavelet-based ECG arrhythmia recognition via sequential particle support vector machine classifiers,” *Journal of Medical and Biological Engineering*, vol. 32, no. 6, pp. 381–396, 2012.
- [248] M. Guillemard, A. Iske, and U. Zölzer, “Clifford Algebras and Dimensionality Reduction for Signal Separation and Classification,” Germany, 2010.

# **Appendix A ECG Signal processing with Wavelet Filter Banks**

## **A.1 Introduction**

There have been significant advances in wavelet theory, over the last ten years. Most of the developments in the theory have involved the development of new bases for various function spaces. These developments include orthonormal wavelets with compact support.

Furthermore, the applications in wavelet transform have been adopted in different fields and industries, such as in the fields of signal processing, image processing and compression (in general). In the field of signal processing, weak signals buried in noise are recoverable through the application of wavelets filters. This is particularly true in medical applications, such as the processing of X-ray images, as well as in magnetic resonance imaging. Images that are processed using this approach can be drawn together without the unnecessary blur, which may muddle the details. In recent decades, several research studies have examined the application of wavelet filter banks in medical signal processing and in particular in the field of electrocardiography.

This chapter provides an overview of wavelet transform (WT) algorithms for analysing ECG records. Aside from the general discussion of WT, this chapter covers continuous wavelet transformation (CWT), discrete wavelet transform (DWT), mother wavelet design, multi-signal wavelet analysis and filter banks. Specific examples on using WT with ECG signals are also provided.

## **A.2 Biomedical signal processing and Electrocardiogram (ECG) signals**

Biomedical signals, such as the heartbeat, have a tendency to be non-stationary; hence, the wavelet transform is probably the most useful tool in analysing heartbeat signals. Through the use of a galvanometer, the electrical potentials connecting a range of body points can be measured. These signals determine whether the heart is functioning normally or possesses a serious health threat or cardiac abnormalities. Different types of filters used for biomedical signal processing, particularly in filtering these noises, include FIR and IIR filters, median

filters, and adaptive filters [1]. High-pass filters are used for the fluctuations of isoelectric lines, often called baselines or isolines. In the FIR filter with a Hamming window, the specifications include a 1500 minimum required filter order, 250 Hz sampling frequency, and 0.35–0.9 Hz transition band. Additionally, a 3-second filter delay is acquired in the FIR filter. The result of the FIR filtration is excellent with the exception of its high delay. In the Butterworth IIR filter, the design parameters only consist of sampling frequency, cut-off frequency, and filter order. Spline interpolation is an alternative method of eliminating the isoelectric line alternations. The estimated isolines can be reconstructed through interpolation. Median filtration, on the other hand, is primarily used for impulse noise suppression. Noise can be regarded as part of the QRS complex and can be eliminated from the ECG signal. Other waves, such as P and T waves, can also be filtered out by repeating the filtration process. A second-order IIR filter and an adaptive filter are commonly used for suppressing power-line noise. Simulink may also be used for adaptive filtration [1]

### A.3 Wavelet transform

Wavelet transform' refers to the decomposition of the quadratic integrable function, which is  $x(t) \in L^2(R)$ , among the family  $\psi(t)$ , whether scaled or translated [2]. WT is similar to the Fourier transform (FT). In that data or functions are approximated with a particular function that has a specific shape. An important difference is that the size of the window used varies, whereas in the Fourier transform remains fixed. According to Liu (2010), the main difference between WT and FT is that FT decomposes the signal into sines and cosines, the functions localised in the Fourier space, whereas WT uses functions that are localised in both the real and the complex Fourier space. The following equation expresses WT [2]:

$$F(a,b) = \int_{-\infty}^{\infty} x(t)\psi_{(a,b)}^*(t)dt \quad (\text{A.1})$$

where\* denotes complex conjugation,  $\psi(t)$  is a function of mother wavelet and  $L^2(R)$  refers to the space of square integrable functions. It should be noted that a wavelet is a small wave that is zero outside its bounded interval. These functions can be real or complex [3].

The frequency components of any signal can be analysed and extracted using an FT. The feature information of time and frequency can be calculated using the short-time Fourier transform (STFT). A Fourier analysis refers to the decomposition of a signal into sine waves

of different frequencies. A Wavelet analysis refers to the decomposition of a signal into shifted and scaled versions of the original wavelet. To provide a useful time-frequency localisation, the WT can be used because it can be reasonably well localised, both in time and frequency. In terms of the scaling function  $\varphi_{j,k}(t)$  and the mother wavelet function  $\psi_{j,k}(t)$ , a signal  $x(t)$  is expressed using wavelet transform as:

$$x(t) = \sum_k a_{j_0,k} \varphi_{j_0,k}(t) + \sum_{j=j_0}^{\infty} \sum_k b_{j,k} \psi_{j,k}(t) \quad (\text{A.2})$$

where  $j$  is the parameter of dilation or the visibility in frequency and  $k$  is the parameter of the position,  $a$  and  $b$  are the coefficients associated with both  $\varphi_{j,k}(t)$  and  $\psi_{j,k}(t)$  respectively.

The coefficients  $a$ ,  $b$  can be calculated using the following equations [4]:

$$a_{j_0,k} = \int_{-\infty}^{\infty} x(t) \varphi_{j_0,k}(t) dt \quad (\text{A.3})$$

$$b_{j_0,k} = \int_{-\infty}^{\infty} x(t) \psi_{j,k}(t) dt \quad (\text{A.4})$$

The scaling function and mother wavelet function can be expressed as:

$$\varphi_{j_0,k}(t) = 2^{j/2} \varphi(2^j t - k) \quad (\text{A.5})$$

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \quad (\text{A.6})$$

The WT can be divided into two main types: continuous (CWT) and discrete (DWT).

### A.3.1 The Continuous Wavelet Transform (CWT)

The CWT is implemented with the utilisation of arbitrary scales and arbitrary wavelets [3]. In a CWT, the wavelets are non-orthogonal. Moreover, there is a high correlation amongst the data or information obtained. CWT is used in the processing of analogue signals and enables their features to be categorised into wavelets [5]. The CWT is used in order to analyse the variation of frequency over time of a signal. Furthermore, common time-varying patterns are revealed by the wavelet coherence of two signals. The analysis of a continuous wavelet is also used in the interpretation of images. The frequency content of an image as it changes across the image can be observed. This helps identify patterns in a noisy image [6].

In the current study, in order to obtain CWTs of images and signals, the MATLAB wavelet toolbox is used. This is often used to obtain the coherence of wavelets between two signals; or to rebuild the time frequency estimates of the signals. Image and signal compression are two of the many applications of the wavelet transform. The WT provides a high image quality and higher compression ratios, as compared to other conventional techniques [7].

The CWT can be applied to areas such as gait analysis, texture analysis, ECG analysis, filter design, detection of transients, detection of corners and edges in an image etc. [8]. The CWT of a function  $x(t)$  is defined as the integral transform of function  $x(t)$ , with a family of wavelet functions (daughter wavelets)  $\psi_{a,b}(t)$ . In order to find the CWT of a function when translated by distance  $b \in R$  at a scale ( $a > 0$ ) the transform can be expressed as:

$$X(a, b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t-b}{a} \right) dt \quad (\text{A.7})$$

This is the most commonly used expression for the CWT.

The wavelet function is designed to strike a balance between time domain and frequency domain resolution. However, it should be noted that the above given equation for WT is a representation of an over-complete function. What should be done is that there should be a reconstruction on the basis of this representation  $x(t)$ . The reconstruction of  $x(t)$  can be obtained using a CWT inversion. The following equation enables the reconstruction of  $x(t)$ , using a CWT inversion.

$$x(t) = \int_0^{\infty} \int_{-\infty}^{\infty} X(a, b) \psi_{a,b} \left( \frac{t-b}{a} \right) da db \quad (\text{A.8})$$

It is notable here that the boundaries regarding translation extend to infinity and those regarding frequency indicate causality. In practical applications the infinity signs on integration are replaced by large constants taking appropriate limits.

### **A.3.2 The discrete wavelet transform**

The DWT is another type of wavelet transform. It should be noted that this is completely different from the previously discussed CWT. The DWT is more commonly used for numerical analysis, signal and image processing [9] while the CWT is more commonly used by physicists to analyse patterns in physical phenomena. Kociolek *et al.* (2001) have

explained that the DWT involves a linear transformation that operates on a data vector, whose length is of an integer power of two, transforming it into a numerically different vector of the same length [10]. DWT is a tool that can be used to separate data into different frequency components. Each component can then be studied with a resolution matched to its scale. It is easily characterised by the multi-scale representation of a function. It is also important to note that the DWT is orthogonal and invertible [11].

Additionally, the DWT is considered the transform of wavelets, when the wavelets are sampled discretely in functional analysis or numeric analysis. It is preferred over other wavelet transforms. Compared to the Fourier transform or the windowed Fourier transform, this also has the advantage of temporal resolution, and has the ability to capture both location information in terms of time and frequency [12].

The practical applications of DWT can be found in the processing of signals in digital communications and when it comes to gait analysis. Generally, the discrete wavelet of discrete signals  $x(t)$  can be expressed through the following equation:

$$x(t) = \sum_k W_\varphi(j_0, k) \varphi_{j_0, k}(t) + \sum_{j=j_0}^{\infty} \sum_k W_\psi(j, k) \psi_{j, k}(t) \quad (\text{A.9})$$

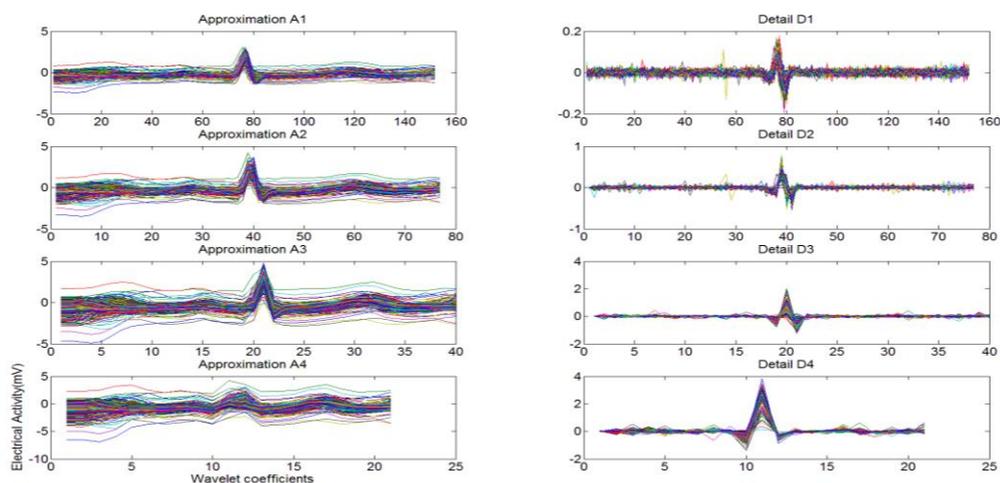
where  $W_\varphi(j_0, k)$  are approximation coefficients,  $W_\psi(j_0, k)$  are detailed coefficients,  $\varphi_{j_0, k}(t)$  scaling function and  $\psi_{j, k}(t)$  is the mother wavelet function.

### A.3.3 Multi-signal Wavelet Analysis

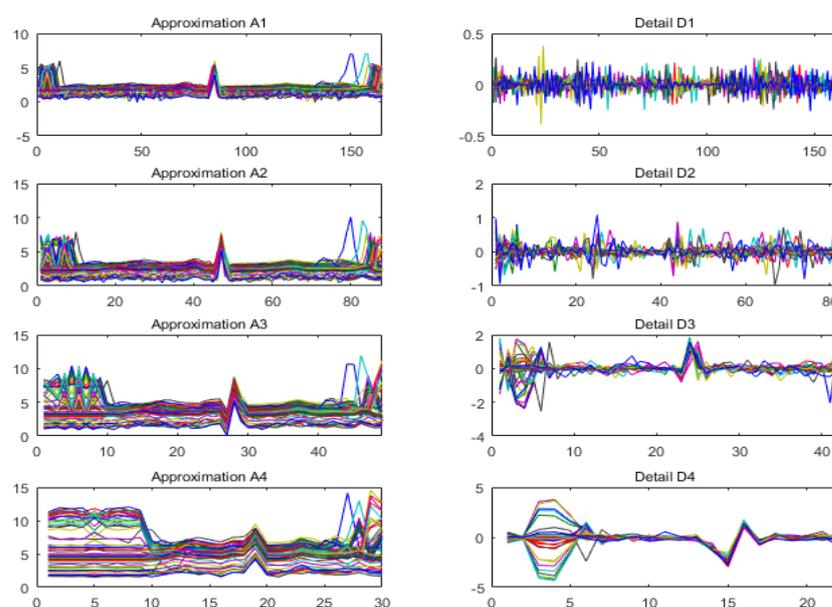
A multi-signal is a set of 1-D signals with the same lengths and located in one matrix, whether in a row or in a column format. The main advantage of using the DWT transform is that it can denoise more than one signal at the same time, before presenting them into the classifier as input. It can also be used with various mother wavelet functions, such as the Daubechies wavelet (db6) and (db4) [13].

Zhao and Yan (2010) have explained that the analysis of multi-signal wavelets has to do with showing how to analyse, denoise, or compress a multi-signal, and then to classify different representations or simplified versions of the signals composing the multi-signal. Multi-signal wavelet analysis allows for the compression of a big set of signals, but only with the minimal utilisation of wavelet representations [14].

It should be noted that there are different levels of resolution for the reconstructed approximations. As stated in [15], denoising and compressing are two of the main applications of wavelets, often used as a pre-processing step before doing classification. Fig A.1 illustrates the use of a multi-signal wavelet decomposition with a normal ECG beat.



**Figure A.1: Analysis Normal ECG beat using multi-signal wavelet decomposition (MWD).**



**Figure A.2: Analysis Normal ECG beat were taken from St Petersburg INCART 12-lead Arrhythmia Database (incartdb) and extracted coefficients using multi-signal wavelet decomposition (MWD)**

### A.3.4 Mother wavelet choices

The mother wavelet involves a transforming function; it gets its name because it is made of smaller waves. Its function is of an oscillatory nature [16].

The multi-resolution analysis at different scales of the mother wavelet also implies that phenomena with fractal (self-repeating) characteristics can be compactly represented. The mother wavelet can be considered a prototype that can generate functions for other windows [16]. The daughter wavelets are translated and scaled copies of the oscillating mother waveform, with fast-decaying finite-length features.

Wavelets are generally derived from a single basic mother wavelet,  $\psi(t)$ . A family of these wavelets can be obtained by scaling and translating  $\psi$ , using the following expression:

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t-\tau}{s}\right) \quad s, \tau \in \mathbb{R}, s \neq 0 \quad (\text{A.10})$$

where  $\psi_{\tau,s}(t)$  is a family of mother wavelet (daughter wavelets),  $s$  is a scaling factor and  $\tau$  is a translation parameter, controlling the location of the wavelet [3].

As discussed by Liu (2010), the mother wavelet function does not extend to infinity in terms of the time domain but is bounded. The  $\psi(t)$  has values in a certain range and zeros elsewhere. Another property of the mother wavelet is that it has zero-mean and that it is normalised. These constraints can be expressed mathematically as follows [2]:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (\text{A.11})$$

$$\|\psi(t)\|^2 = \int_{-\infty}^{\infty} \psi(t) \psi^*(t) dt = 1 \quad (\text{A.12})$$

There are different types of mother wavelets, such as Haar, Daubechies, Biorthogonal, Coiflets, Symlets and Morelet, as will be discussed in the following section.

#### A.3.4.1 The Daubechies wavelets db(N)

The Daubechies wavelets db(N) were defined by Ingrid Daubechies in 1992 and they are called compactly supported orthonormal wavelets. The families of these wavelets are written as dbN where N is the order and db is the family name of the wavelet. The Daubechies wavelet function has been used to decompose data as shown in Fig A.3-A.6, which uses Daubechies (db4). Daubechies is a type of wavelet that can find the minimum size discrete

filter and for a given vanishing moment  $p$ . These properties of wavelets also have orthogonal and compact support.

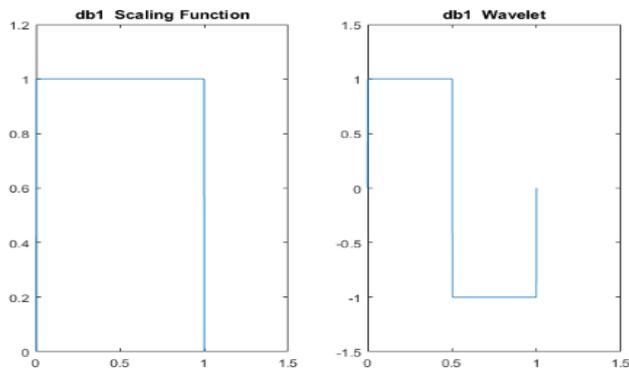


Figure A.3: db1 Scaling and wavelet Function.

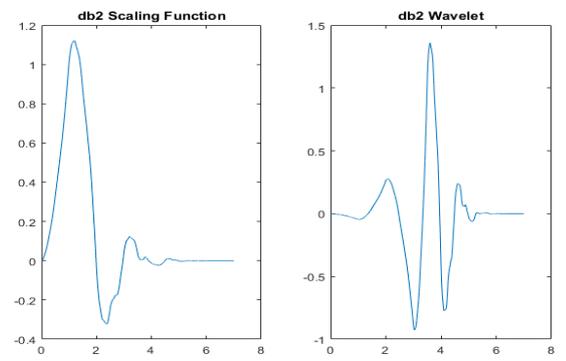


Figure A.4: db2 Scaling and wavelet Function.

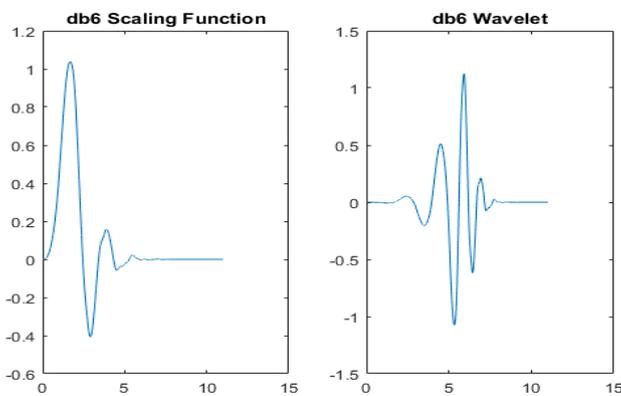


Figure A.5: db4 Scaling and wavelet Function.

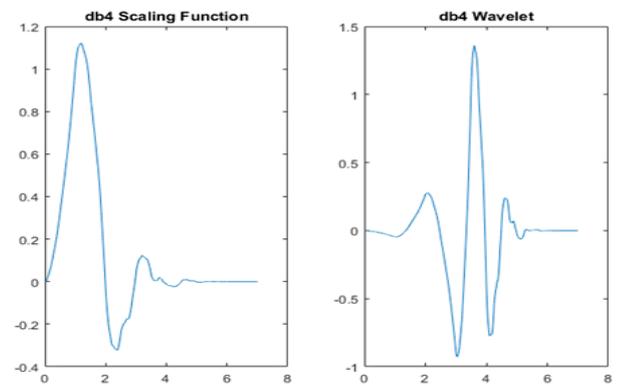


Figure A.6: db6 Scaling and wavelet Function.

### A.3.4.2 The Symlet wavelet (symN)

The symN wavelets are also known as Daubechies least-asymmetric wavelets. The symlets are more symmetric than the extremal phase wavelets. The  $N$  represents the number of vanishing moments in symN, and the symlets properties include orthogonal and compact support. The filter length is  $2p$ , it has  $p$  vanishing moments and is nearly linear in phase.

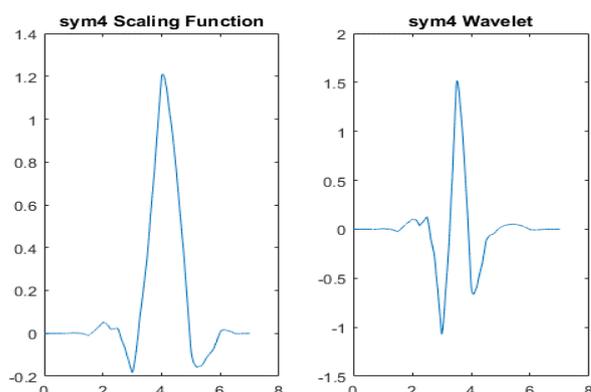


Figure A.7: sym4 scaling and wavelet function.

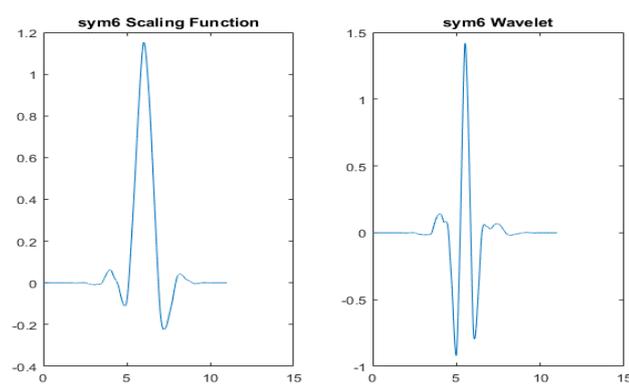


Figure A.8: sym6 scaling and wavelet function.

### A.3.4.3 The Coiflet Wavelets (coifN)

This function can be used in applications in numerical analysis. Coifman wavelets are constructed as a family of wavelets that have  $p$  vanishing moments with some additional requirement about the vanishing moments of the scaling function. The equations (A.13) and (A.14) explain the two requirements, regarding vanishing moments of the scaling function. In  $\text{coifN}$ , the number of vanishing moments for both the wavelets and scaling functions is  $N$ . For the coiflet construction, the number of filter coefficients is  $3N$ . General characteristics of the Coiflets Wavelets are compactly supported wavelets with the highest number of vanishing moments. Fig A.9 shows  $\text{coif5}$  scaling and wavelet Function.

$$\int_{-\infty}^{\infty} \phi(t) dt = 1 \tag{A.13}$$

$$\int_{-\infty}^{\infty} t^k \phi(t) dt = 0 \quad \text{for } 1 \leq k < p \tag{A.14}$$

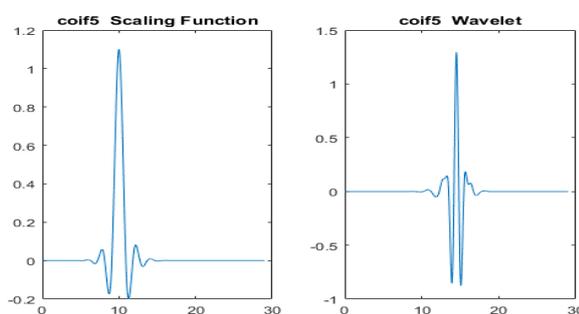


Figure A.9: coif5 scaling and wavelet Function.

#### A.3.4.4 The Haar wavelet (haar)

The Haar wavelet is the oldest and the simplest wavelet, which has a minimum vanishing moment. The scaling function is symmetric while the wavelet function is anti-symmetric. It represents the same wavelet as db1 as shown in Fig A.3.

### A.4 Wavelets from a filter banks perspective

The filter banks, such as the biorthogonal and orthogonal, are considered the arrangements of bandpass, high pass and low pass filters, that can divide the data into sub bands. If the sub bands are not modified, the perfect reconstruction of the original data is possible, by using inverse filters. In most of the applications, the data is processed differently in the different sub bands. This enables the reconstruction of a modified version of the original data. A linear phase does not occur in the orthogonal filter banks, but there is a linear phase in the biorthogonal filter banks [17]. This implies that biorthogonal filters are well placed; for example, in terms of the analysis and synthesis of a control system.

The scaling function can be specified by the number of vanishing moments in the wavelet function. This helps retain or remove the polynomial behaviour in the data. The wavelet toolbox function can be used to obtain the most common orthogonal and biorthogonal wavelet filters [18]. The perfect reconstruction filters can be designed by the elementary steps of lifting. In the reconstruction phase, the original signal is synthesised from the coefficients that are obtained at each level of the decomposition stages, as shown in Figure A.10. This transform is also known as the ‘Inverse Discrete Wavelet Transform’ (IDWT). This technique is most commonly used in sub-band coding. Herley and Vetterli (1999) further explained: “The wavelet transform has recently emerged as a powerful tool for non-stationary signal analysis. Its discrete version is closely related to filter banks which have been studied in digital signal processing”. A biorthogonal filter bank can be constructed with the use of a linear phase FIR filters.

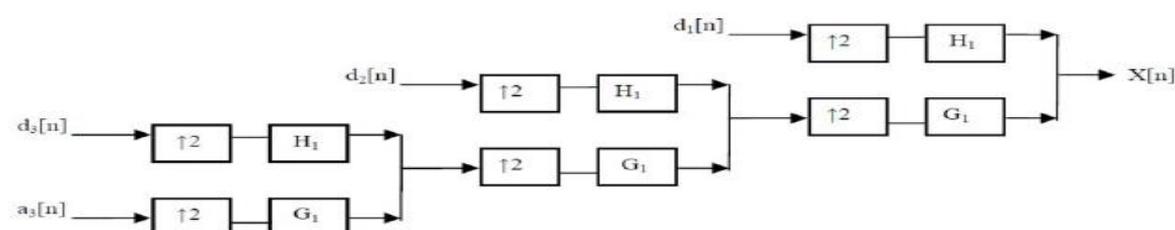


Figure A.10: Filter bank for perfect reconstruction.

### **A.4.1 Wavelet filter bank signal processing of biomedical signals**

Various types of filters were considered used in order to analyse the ECG signals or detect the beats. For example a nonlinear filter technique and first-order derivative moving average filters have been used for QRS complex detection and heart-rate calculation [19]. In 2009, Balasubramaniam and Nedumaran indicated that ECG signal processing can be seen as a two-stages preliminary processing and primary processing routine. In the first stages, artefacts such as higher peaks due to electrode motion and power-line 60 Hz interference are removed, while in primary processing, techniques such as denoising, baseline wandering and detection of P, QRS, and T waveforms are achieved through the use of some of the algorithms [20].

As was mentioned, ECG waveforms consist of five basic waves P, Q, R, S, and T waves and, sometimes, U waves. Under normal conditions, cardiac cells are electrically polarised and a normal heartbeat has a usual ECG tracing that is composed of a P wave that causes a contraction of the atrium, a QRS complex that causes left and right ventricular contractions, and a T wave that signifies ventricular repolarisation. The major disadvantage of using as FFT to analyze these waves is its inability to provide the exact location information of the frequency components. The short-term Fourier transform (STFT), may also be used but its drawback is that the precision of both time and frequency is not optimal. Hence, wavelet transformation is a more suitable approach for the studying of ECG signals.

### **A.4.2 Wavelets filter banks and their use in Electrocardiogram (ECG) signals analysis.**

Multi-rate digital filters and filter banks are often applied in the digital audio industry, communications, antenna systems, speech processing, and image compression [21]. A multi-rate digital signal processing algorithm has been designed in this study in order to perform multiple tasks in ECG processing. This algorithm includes the use of filter banks to decompose the ECG signals into sub-bands with identical frequency bandwidths. A filter bank (FB) is a set of band pass filters which break up the input signal into multiple components, each of which possesses the original signal's single frequency sub-band.

Performing signal processing by means of a bank of digital filters in combination with down-sampling is the idea behind filter bank theory. This process is based on a given signal  $x$

passing through a low pass filter with a transfer function  $C(z)$  and a high pass filter with a transfer function  $D(z)$  at each stages of the filtering process.

In addition, these two filters have to meet certain requirements in order to enable a perfect reconstruction of the two output signals that are obtained after down-sampling a corresponding mother wavelet  $\psi(t)$ . Having compact support leads to an improvement in the efficiency of the decomposition and reconstruction. Let us assume two polynomials in the  $\mathbb{Z}$  domain:

$$C(z) = c_0 + c_1 z^{(-1)} + c_2 z^{(-2)} + \dots + c_N z^{(-N)} \quad (\text{A.15})$$

$$D(z) = d_0 + d_1 z^{(-1)} + d_2 z^{(-2)} + \dots + d_N z^{(-N)} \quad (\text{A.16})$$

where  $N$  is an integer determining the order of the filters. The orthogonality of the underlying wavelets on the filter coefficients includes the following conditions:

Normalization:

$$\sum_{k=0}^N C_k^2 = 1 \quad \text{and} \quad \sum_{k=0}^N d_k^2 = 1 \quad (\text{A.17})$$

Double shift orthogonality

$$\sum_{k=0}^N C_k C_{k-2l} = 0 \quad \text{and} \quad \sum_{k=0}^N d_k d_{k-2l} = 0 \quad (\text{A.18})$$

for all integers  $l \neq 0$

(c) *double shift orthogonality between two filters:*

$$\sum_{k=0}^N C_k d_{k-2l} = 0 \quad (\text{A.19})$$

for all integers  $l$ .

A double shift orthogonality implies that  $N$  is odd; for example,  $N=2n-1$ . [22]. Analysis and synthesis are the two basic types of filter banks. An analysis bank uses filters  $H_l(z)$ , which split a signal into  $M$  subband signals  $x_l(n)$ , as shown in Figure A.2. A synthesis bank consists

of  $M$  synthesis filters  $F_l(z)$ , which combine  $M$  signals  $y_l(n)$  into a reconstructed signal. Very often by marking some of these coefficients equal to zero, we actually filter the signal.

The analysis filters are based on decomposing an incoming signal into specific frequency bands or sub-bands, while synthesis filters use the processed sub-bands, resulting in a processed version of the input signal. An FB-based algorithm involves decomposing a signal into frequency sub-bands and then reconstructing the processed sub-bands [21].

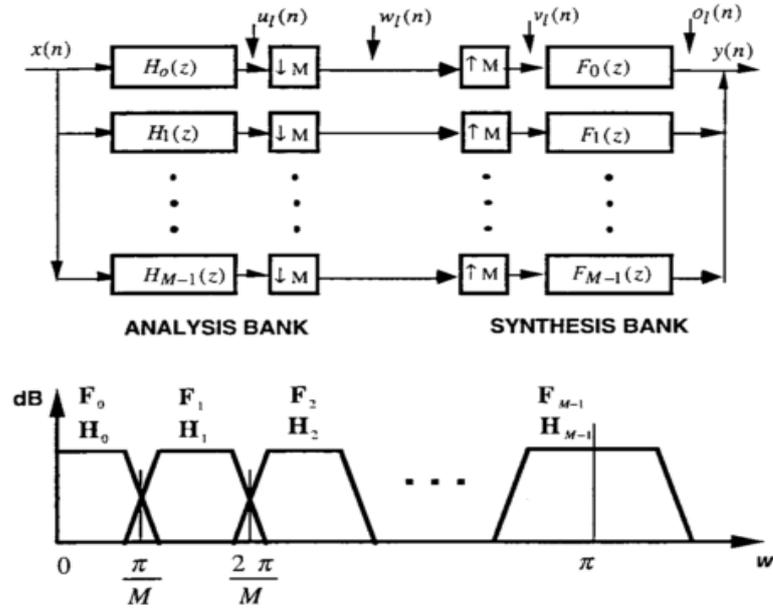


Figure A.11: Analysis and synthesis filter banks adopted from based on single scale multi-resolution analysis, adopted from [23].

As can be seen in Figure A.11 an filter bank involves  $M$  analysis and  $M$  synthesis filters, while  $L$  is the length of each sub-band signal  $U_l(z)$ ,  $l = 0, 1, \dots, M-1$ , found by using the analysis filters as the band passes the input signal  $X(z)$ .

$$H_l(z), l = 0, 1, \dots, M-1 \quad (\text{A.20})$$

$$U_l(z) = H_l(z) X(z) \quad l = 0, 1, \dots, M-1 \quad (\text{A.21})$$

As  $\pi/M$  is the effective bandwidth of  $U_l(z)$ , it can be down-sampled to decrease the total rate by using the following equation

$$W_l(z) = \frac{1}{M} \sum_{k=0}^{M-1} U_l\left(z^{1/M} W^k\right) \quad l = 0, 1, \dots, M-1 \quad (\text{A.22})$$

where  $W_l(z)$  is the down-sampled signal,  $W = e^{-j(2\pi/M)}$ .

The up-sampling operation is based on inserting zeros after each point in the time domain sequence discretized time domain.

$$V_l(z) = W_l(z^M) = \frac{1}{M} \sum_{k=0}^{M-1} H_l(zW^k) X(zW^k) \quad l = 0, 1, \dots, M-1 \quad (\text{A.23})$$

The synthesis filters can be operated efficiently by using the following expression

$$O_l(z) = F_l(z)V_l(z) \quad l = 0, 1, \dots, M-1 \quad (\text{A.24})$$

where  $F_l(z)$  represents the synthesis filters and  $F_l(z)$  the sub-band signals  $O_l(z)$ . The sub-band signals  $O_l(z)$  can be algebraically added point by point to result in the output  $Y(z)$  [23].

$$Y(z) = \sum_{l=1}^{M-1} O_l(z) \quad (\text{A.25})$$

$$= \sum_{k=0}^{M-1} F_l(z)V_l(z) \quad (\text{A.26})$$

$$= \frac{1}{M} \sum_{k=0}^{M-1} X(zW^k) \sum_{k=0}^{M-1} H_l(zW^k) F_l(z) \quad (\text{A.27})$$

Furthermore, some of the ECG processes which can be performed using filter banks include detection of heartbeat, beat classification, noise alert, and ECG enhancement [24]. In ECG processing, the analysis filters are required to have a linear phase and identical group delay since the fiducial points across the sub-bands must have the same delay. In the algorithm of beat detection, the presence of a beat is indicated by the energy in the region of the QRS filter. Here, the SNR (signal-to-noise ratio) of the QRS complex is maximized by the band pass filter, and then the band pass signal is squared and entered into a moving window integrator (MWI) to obtain an approximation of its energy [23]. Some result of using DWT and filter bank are discussed in the following section.

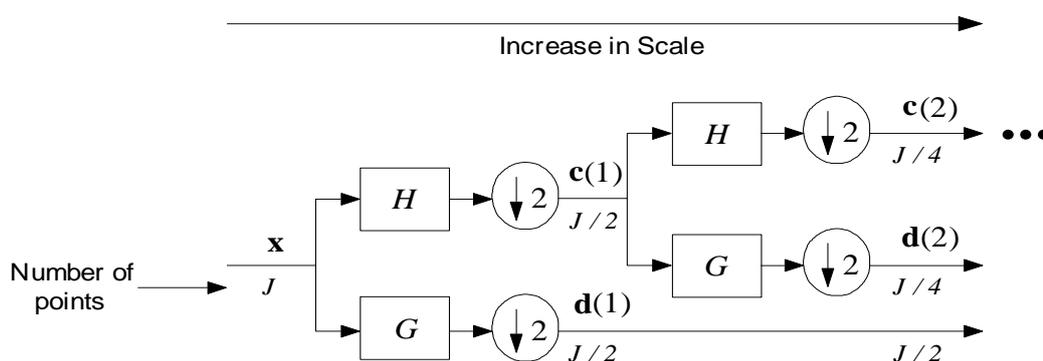
Additionally, ECG enhancement can be made possible through the processing of the sub-bands, which is dependent on the time and frequency, and then reconstructing them. Some results of using DWT and FB will be mentioned in this Chapter.

#### A.4.2.1 ECG signal decomposition using Discrete Wavelet Transforms (DWTs)

In the signal decomposition using the DWT, both a low pass filter bank (LPF) and a high pass filter bank (HPF) are used to generate time domain responses, these are convolved with the time domain ECG signal. Convolution of the response function of the chosen filter

(corresponding to a particular mother wavelet) with the signal provides an output which has different energy at different scales. Approximation coefficients relate to the low frequency components of the signal whereas detail coefficients relate to the higher frequency components in the signal. Wavelet decomposition using the DWT provides essentially a multi-resolution representation of the input signal. The user normally retains coefficients up to a particular scale whereas more detailed decompositions become redundant as their incorporation have a negligible effect on the signal. The convolution operation may be conveniently performed in the frequency domain where it is implemented through a simple multiplication process [25].

The ECG signals considered in this study were taken from the European ST-T Database. The features in the ECG signal were extracted using DWTs from the sym3, db4 and db6 wavelet families. Six decomposition levels are more than sufficient to faithfully represent the ECG signals. As shown clearly in Fig A.12, blocks  $H[n]$  and  $G[n]$  represent the low-pass and high-pass filter responses ( $i=1, \dots, 6$ ) respectively, and the  $\downarrow 2$  operator denotes dyadic down-sampling. Approximation  $c_i$  and Detail  $d_i$  coefficients at each decomposition step are also shown. Only perfect reconstruction quadrature mirror filter banks (orthogonal transforms) are considered in this study because they fully preserve the information content in the signal. This is important from an algorithm certification perspective which is normally associated with the introduction of new algorithms for bio-medical software applications. The discrete wavelet transform can be calculated in a fast manner by using finite-impulse-response (FIR) filter banks.



**Figure A.12:** Two-channel filter bank implementation of the wavelet transform applied to a data vector  $x$ . Blocks  $H$  and  $G$  represent a lowpass and a highpass filter respectively and  $\downarrow 2$  denotes the operation of dyadic downsampling. The decomposition can be carried out in more resolution levels by successively splitting the lowpass channel, adopted from [26].

The filter bank transform can be regarded as a change of variables for the ECG signals  $x_n$ :

$$t_j = \sum_{n=0}^{J-1} x_n v_j(n), \quad j = 0, 1, \dots, J-1 \quad (\text{A.28})$$

where  $t_j$  is a transformed variable and  $v_j(n) \in \mathfrak{R}$  is a transform weight. The transfer function of the low-pass filter in the z-domain can be written as:

$$H^{(N)}(z) = \sum_{n=0}^{2N-1} h_n^{(N)} z^{-n} = H_0^{(N)}(z^2) + z^{-1} H_1^{(N)}(z^2) \quad (\text{A.29})$$

where superscript  $(N)$  denotes that the filtering sequences have length  $2N$  and  $H_0^{(N)}(z)$  and  $H_1^{(N)}(z)$  denote polyphasic components of  $H^{(N)}(z)$ . Further algorithmic details of the proposed filter banks and a discussion of adaptive filter banks for the purpose are discussed in [26]. Results from multi-level decompositions are shown in Figure A.13.

Of particular interest to the current study is to identify the most parsimonious representation of the ECG signals in the wavelet domain so that a non-linear neural network classifier can perform the classification task directly in the wavelet domain. Further parameterization of the signals using adaptive wavelets using various adaptive filter banks [26] from the wavelet transform literature are considered in this project. This example which was also published in [27] shows the result of using multilevel decomposition of an ECG signal with some wavelet families such as sym3, db4 and db6.

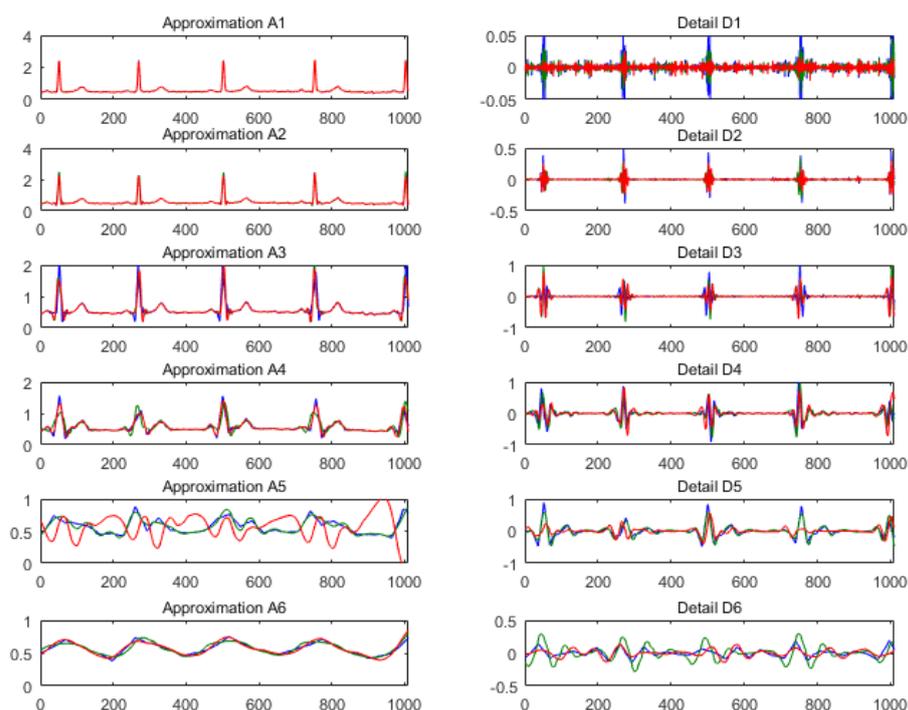


Figure A.13: ECG signal multilevel decomposition using sym3, db4 and db6 wavelet families.

### A.4.3 An optimal wavelet filter banks for ECG signal classification

Signal decomposition using perfect reconstruction quadrature mirror filter banks can provide a very parsimonious representation of ECG signals. In a previous work [26], it was shown that optimal wavelets can be used for the post-processing of ECG signals so that classifiers can operate directly in the wavelet domain as opposed to the time or frequency domains. This section will discuss the extension of the wavelet parametrization approach proposed by Sherlock and Monroe [28] to ensure that the derived wavelets have at least two vanishing moments.

### A.4.4 Wavelet filter bank parametrisation

In the signal decomposition using the DWT, both a low pass (LPF) and a high pass (HPF) filter bank are used to generate time domain responses, these are convolved with the time domain ECG signal. Convolution of the response function of the chosen filter (corresponding to a particular mother wavelet) with the signal provides an output which has different energy at different scales. Approximation coefficients relate to the low frequency components of the signal whereas detail coefficients relate to the higher frequency components in the signal.

Wavelet decomposition using the DWT provides essentially a multi-resolution representation of the input signal. The user normally retains coefficients up to a particular scale whereas more detailed decompositions become redundant as their incorporation have a negligible effect on the signal. The convolution operation may be conveniently performed in the frequency domain where it is implemented through a simple multiplication process.

In this filter bank, the low-pass filtering result undergoes successive filtering iterations with the number of iterations  $N_{it}$  chosen by the analyst. The final result of the decomposition of data vector  $\mathbf{x}$  is a vector resulting from the concatenation of row vectors  $\mathbf{c}(N_{it})$  (termed approximation coefficient at the largest scale level) and  $\mathbf{d}(s)$  (termed detail coefficients at the  $s^{\text{th}}$  scale level,  $s = 1, \dots, N_{it}$ ) in the following manner:

$$\mathbf{t} = [\mathbf{c}(N_{it}) \mid \mathbf{d}(N_{it}) \mid \mathbf{d}(N_{it} - 1) \mid \dots \mid \mathbf{d}(1)] \quad (\text{A.30})$$

with coefficients in larger scales (e.g.  $\mathbf{d}(N_{it}), \mathbf{d}(N_{it} - 1), \mathbf{d}(N_{it} - 2), \dots$ ) associated with broad features in the data vector, and coefficients in smaller scales (e.g.  $\mathbf{d}(1), \mathbf{d}(2), \mathbf{d}(3), \dots$ ) associated with narrower features such as sharp peaks. The filter bank transform can be regarded as a change in variables from  $\mathfrak{R}^J$  to  $\mathfrak{R}^J$  performed according to the following operation,

$$t_j = \sum_{n=0}^{J-1} x_n v_j(n), \quad j = 0, 1, \dots, J-1 \quad (\text{A.31})$$

where  $t_j$  is a transformed variable and  $v_j(n) \in \mathfrak{R}$  is a transform weight. It proves convenient to write the transform in matrix form as:

$$\mathbf{t}_{1 \times J} = \mathbf{x}_{1 \times J} \mathbf{V}_{J \times J} \quad (\text{A.32})$$

where  $\mathbf{x} = [x_0 \ x_1 \ \dots \ x_{J-1}]$  is the row vector of original variables,  $\mathbf{t}$  is the row vector of new (transformed) variables and  $\mathbf{V}$  is the matrix of weights. Choosing  $\mathbf{V}$  to be unitary (that is,  $\mathbf{V}^T \mathbf{V} = \mathbf{I}$ ), the transform is said to be orthogonal and it, therefore, consists of a simple rotation in the coordinate axes (with the new axes directions determined by the columns of  $\mathbf{V}$ ).

Let  $\{h_0, h_1, \dots, h_{2N-1}\}$  and  $\{g_0, g_1, \dots, g_{2N-1}\}$  be the impulse responses of the low-pass and high-pass filters respectively. Assuming that filtering is carried out by circular convolution, the procedure for generating the approximation coefficients from the data vector  $\mathbf{x}$  is illustrated in Table A.1. The convolution consists of flipping the filtering sequence and moving it alongside the data vector. For each position of the filtering sequence with respect to

the data vector, the scalar product of the two is calculated (with missing points in the filtering sequence replaced with zeros). For instance, if  $N = 2$ , the third row in Table 1 shows that  $c_1' = x_1h_3 + x_2h_2 + x_3h_1 + x_4h_0$ . Dyadic down-sampling is then performed to  $c_{2i}'$  to generate coefficients  $c_i$ . The detail coefficients  $d_i$  are obtained in a similar manner by using the high-pass filtering sequence.

**Table A.1.** Convolution procedure for low-pass filtering showing results before and after dyadic down-sampling.

$x_0$	$x_1$	$\dots$	$x_{2N-1}$	$x_{2N}$	$\dots$	$x_{J/2-1}$	$x_0$	$x_1$	$\dots$	$x_{2N-2}$	Before	After
$h_{2N-1}$	$h_{2N-2}$	$\dots$	$h_0$				$\dots$				$c_0'$	
	$h_{2N-1}$	$\dots$	$h_1$	$h_0$			$\dots$				$c_1'$	$c_0$
		$\dots$	$\vdots$				$\dots$				$\vdots$	$\vdots$
						$h_{2N-1}$	$h_{2N-2}$	$h_{2N-3}$	$\dots$		$c_{J/2}'$	
						$h_{2N-1}$	$h_{2N-2}$	$\dots$	$h_0$	$c_{J-1}'$	$c_{J/2-1}$	

If the approximation  $\mathbf{c}$  and detail  $\mathbf{d}$  coefficients are stacked in vector  $\mathbf{t} = [\mathbf{c} \mid \mathbf{d}]$ , the wavelet transform can be expressed in the matrix form with the transformation matrix given by:

$$\mathbf{V} = \begin{bmatrix} 0 & 0 & \dots & h_{2N-4} & h_{2N-2} & 0 & 0 & \dots & g_{2N-4} & g_{2N-2} \\ h_{2N-1} & 0 & \dots & h_{2N-5} & h_{2N-3} & g_{2N-1} & 0 & \dots & g_{2N-5} & g_{2N-3} \\ h_{2N-2} & 0 & \dots & h_{2N-6} & h_{2N-4} & g_{2N-2} & 0 & \dots & g_{2N-6} & g_{2N-4} \\ h_{2N-3} & h_{2N-1} & \dots & h_{2N-7} & h_{2N-5} & g_{2N-3} & g_{2N-1} & \dots & g_{2N-7} & g_{2N-5} \\ \vdots & \vdots \\ h_0 & h_2 & \dots & 0 & 0 & g_0 & g_2 & \dots & 0 & 0 \\ 0 & h_1 & \dots & 0 & 0 & 0 & g_1 & \dots & 0 & 0 \\ 0 & h_0 & \dots & 0 & 0 & 0 & g_0 & \dots & 0 & 0 \\ \vdots & \vdots \\ 0 & 0 & \dots & h_{2N-2} & 0 & 0 & 0 & \dots & g_{2N-2} & 0 \\ 0 & 0 & \dots & h_{2N-3} & h_{2N-1} & 0 & 0 & \dots & g_{2N-3} & g_{2N-1} \end{bmatrix} \quad (\text{A.33})$$

A requirement for the transform to be orthogonal (i.e.,  $\mathbf{V}^T\mathbf{V}=\mathbf{I}$ ) is that the sum of the squares of each column must be equal to one and the scalar product of different columns must be equal to zero [29]. Therefore, for a filter bank that utilizes low-pass and high-pass filters, the following conditions ensure orthogonality of the transform so that no information is lost in the decomposition process [30]:

$$\sum_{n=0}^{2N-1-2l} h_n h_{n+2l} = \begin{cases} 1, & l=0 \\ 0, & 0 < l < N \end{cases} \quad (\text{A.34a})$$

$$g_n = (-1)^{n+1} h_{2N-1-n}, \quad n=0, 1, \dots, 2N-1 \quad (\text{A.34b})$$

Under these conditions, the filter bank is said to enjoy a perfect reconstruction (PR) property, because  $\mathbf{x}$  can be reconstructed from  $\mathbf{t}$ , which means that there is no loss of information in the decomposition process. Although other non-orthogonal filter bank transforms can also enjoy a PR property, provided that they are associated to a non-singular matrix  $\mathbf{V}$ , the analysis in the present work is restricted to orthogonal transforms. In fact, the orthogonality of the transform (with the consequent PR property) ensures that no information that may be potentially useful for classification purposes is lost in the decomposition process. Moreover, convenient parameterisation schemes may then be employed to cast the transform filters into forms amenable to optimization.

The parameterisation of PR FIR filter banks proposed by Vaidyanathan (1993) [29] as adapted by Sherlock and Monro (1998) [28] to parameterise orthonormal wavelets of arbitrary compact support may be used for this purpose. For a filter bank of the form shown in Fig. A.12 where the conditions in Equations (A.34a) and (A.34b) are satisfied, the transfer function of the low-pass filter in the  $z$ -domain can be written as:

$$H^{(N)}(z) = \sum_{n=0}^{2N-1} h_n^{(N)} z^{-n} = H_0^{(N)}(z^2) + z^{-1} H_1^{(N)}(z^2) \quad (\text{A.35})$$

where superscript  $(N)$  denotes that the filtering sequences have length  $2N$ . The terms  $H_0^{(N)}(z)$  and  $H_1^{(N)}(z)$ , which denote polyphasic components of  $H^{(N)}(z)$ , are given by:

$$H_0^{(N)}(z) = \sum_{i=0}^{N-1} h_{2i}^{(N)} z^{-i} \quad (\text{A.36a})$$

$$H_1^{(N)}(z) = \sum_{i=0}^{N-1} h_{2i+1}^{(N)} z^{-i} \quad (\text{A.36b})$$

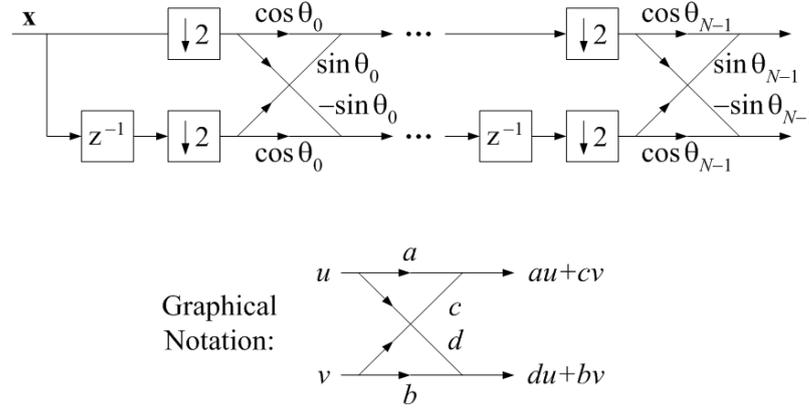
Defining the polyphasic components  $G_0^{(N)}(z)$  and  $G_1^{(N)}(z)$  of the high-pass filter  $G^{(N)}(z)$  in a similar manner, a matrix  $F^{(N)}(z)$  may be defined:

$$F^{(N)}(z) = \begin{bmatrix} H_0^{(N)}(z) & H_1^{(N)}(z) \\ G_0^{(N)}(z) & G_1^{(N)}(z) \end{bmatrix} \quad (\text{A.37})$$

It can be shown [28],[31] that  $F^{(N)}(z)$  can be factorized as:

$$F^{(N)}(z) = \begin{bmatrix} C_0 & S_0 \\ -S_0 & C_0 \end{bmatrix} \prod_{k=1}^{N-1} \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \begin{bmatrix} C_k & S_k \\ -S_k & C_k \end{bmatrix} \quad (\text{A.38})$$

where each pair of parameters  $(C_k, S_k)$  are related to a common angular parameter  $\theta_k$  as  $C_k = \cos(\theta_k)$  and  $S_k = \sin(\theta_k)$ ,  $k = 0, 1, \dots, N-1$ . It follows that the filters can be completely parameterised by  $N$  angles  $\theta_0, \theta_1, \dots, \theta_{N-1}$ , which can assume any value in the set of real numbers, as shown in Fig 4.14.



**Figure: A.14** Procedure for parameterizing wavelet filter banks by  $N$  angles. (a)  $S_k$  and  $C_k$  represent the sine and cosine of angular parameter  $\theta_k$ , respectively, adopted from [26]

The weights of the low-pass filter can be easily recovered from a set of angles  $\{\theta_k\}$  by using the following recursive formula [29]:

$$F^{(k+1)}(z) = F^{(k)}(z) \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} \begin{bmatrix} C_k & S_k \\ -S_k & C_k \end{bmatrix} \quad (\text{A.39})$$

for  $k = 1, 2, \dots, N-1$  with

$$F^{(1)}(z) = \begin{bmatrix} C_0 & S_0 \\ -S_0 & C_0 \end{bmatrix}. \quad (\text{A.40})$$

Equation (A.39) with the initial condition of Equation (A.40) provides a way to obtain the weights  $\{h_i^{(k+1)}\}$  for a filter of length  $2(k+1)$  from the weights  $\{h_i^{(k)}\}$  for a filter of length  $2k$ . To do that, one starts by writing, from Equations (A.37) and (A.39),

$$H_0^{(k+1)}(z) = H_0^{(k)}(z)C_k - z^{-1}H_1^{(k)}(z)S_k, \quad k = 1, 2, \dots, N-1 \quad (\text{A.41a})$$

$$H_1^{(k+1)}(z) = H_0^{(k)}(z)S_k + z^{-1}H_1^{(k)}(z)C_k, \quad k = 1, 2, \dots, N-1 \quad (\text{A.41b})$$

with  $H_0^{(1)}(z) = C_0$  and  $H_1^{(1)}(z) = S_0$ . Then, a recursive formula for the generation of low-pass filter weights with even indexes  $\{h_{2i}\}$  can be stated by using the definitions in Equations (A.36a) and (A.36b) to expand Equation (A.41a) as

$$\underbrace{\sum_{i=0}^k h_{2i}^{(k+1)} z^{-i}}_{H_0^{(k+1)}(z)} = \underbrace{\left( \sum_{i=0}^{(k-1)} h_{2i}^{(k)} z^{-i} \right)}_{H_0^{(k)}(z)} C_k - z^{-1} \underbrace{\left( \sum_{i=0}^{(k-1)} h_{2i+1}^{(k)} z^{-i} \right)}_{H_1^{(k)}(z)} S_k \Rightarrow \quad (\text{A.42})$$

$$\sum_{i=0}^k h_{2i}^{(k+1)} z^{-i} = C_k h_0^{(k)} + \sum_{i=1}^{k-1} (C_k h_{2i}^{(k)} - S_k h_{2i-1}^{(k)}) z^{-i} - S_k h_{2k-1}^{(k)} z^{-k}$$

for  $k = 1, 2, \dots, N-1$ , with  $h_0^{(1)} = C_0$  and  $h_1^{(1)} = S_0$ . From the identity of terms with the same power of  $z$  in the last line of Eq. (A.42), it follows that:

$$\begin{cases} h_0^{(k+1)} = C_k h_0^{(k)} \\ h_{2i}^{(k+1)} = C_k h_{2i}^{(k)} - S_k h_{2i-1}^{(k)}, \quad i = 1, 2, \dots, k-1 \\ h_{2k}^{(k+1)} = -S_k h_{2k-1}^{(k)} \end{cases} \quad (\text{A.43a})$$

for  $k = 1, 2, \dots, N-1$ .

A similar formula can be stated for the low-pass filter weights with odd indexes, by expanding Equation (A.41b) as:

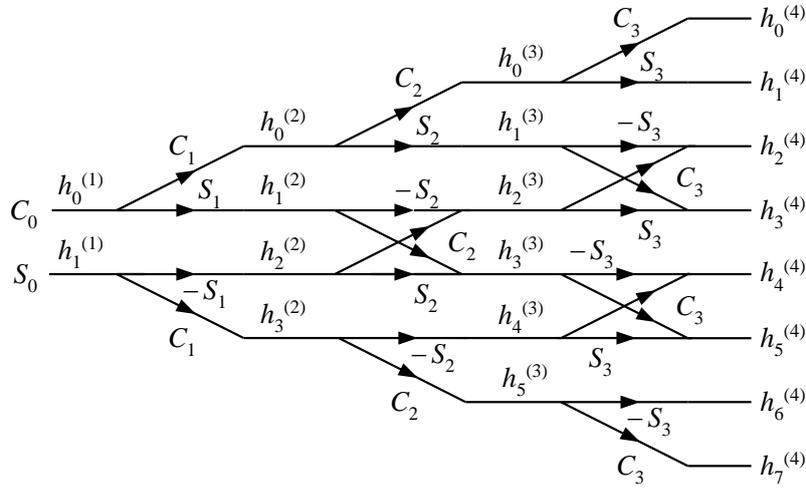
$$\begin{cases} h_1^{(k+1)} = S_k h_0^{(k)} \\ h_{2i+1}^{(k+1)} = S_k h_{2i}^{(k)} + C_k h_{2i-1}^{(k)}, \quad i = 1, 2, \dots, k-1 \\ h_{2k+1}^{(k+1)} = C_k h_{2k}^{(k)} \end{cases} \quad (\text{A.43b})$$

for  $k = 1, 2, \dots, N-1$ . After obtaining the low-pass filtering sequence as explained above, the high-pass filtering sequence can be obtained by using Eq.(A.34b).

The procedure for obtaining filter weights  $\{h_i^{(k+1)}\}$  in terms of  $\{h_i^{(k)}\}$  by adding one additional angular parameter  $\theta_k$  is depicted in Figure A.15 [28]. In this graphical notation, each arrow represents the multiplication of the element at the base of the arrow with the

constant on top of the arrow. When two arrows arrive at the same point, the results of the multiplications are added together.

By adopting the parameterisation described above, the adjustment of the filter bank to the ensemble of signals under consideration can be formulated as a problem of unconstrained optimisation in  $\mathbb{R}^N$ . The optimal filtering procedure employed in this work was aimed at maximising the variance explained by the wavelet coefficients kept in the thresholding process. The optimisation consisted of maximising an objective function  $F(\theta): \mathbb{R}^N \rightarrow \mathbb{R}$  defined as



**Figure A.15:** Recursive generation of lowpass filter weights  $\{h_n^{(k+1)}\}$  in terms of  $\{h_n^{(k)}\}$  by adding one additional angular parameter at a time as discussed by Sherlock and Monro. By using this algorithm, any set of  $N$  angles  $\{\theta_0, \theta_1, \dots, \theta_{N-1}\}$  leads to a sequence of lowpass filter weights that satisfies the orthogonality condition were adopted from (A.32), adopted from [26].

$$F(\theta) = \sum_{k \in I} \sigma^2(j; \theta) \tag{A.44}$$

where  $\theta$  is the vector of  $N$  angles that parameterise the filter bank as explained above,  $\sigma(j; \theta)$  is the standard deviation of the  $j^{\text{th}}$  wavelet coefficient calculated in the set of training signals, and  $I$  is the index set of the coefficients used. Since the overall variance of the data set is preserved by an orthogonal transform, maximizing (A.44) amounts to maximizing the relative explained variance of the wavelet coefficients employed. It is worth noting that  $I$  is defined on the basis of the variance of the wavelet coefficients before the optimisation.

An approach to circumvent the local maxima problem of the formulation described above consists of exploiting the product filter parameterisation proposed by Moulin *et al.*, [32]. In order to describe the optimisation process, we use again a circular convolution process for the data vectors under the down sampling operation  $\mathbf{c}' = \mathbf{h}\tilde{\mathbf{X}}$  and  $\mathbf{d}' = \mathbf{g}\tilde{\mathbf{X}}$  where  $h = [h_{2N-1}, h_{2N-2}, \dots, h_0]$  and  $g = [g_{2N-1}, g_{2N-2}, \dots, g_0]$  are the corresponding impulse response sequences of the low- and high-pass filters respectively and  $\tilde{\mathbf{X}}$  is a circulant matrix formed from the data vector  $\mathbf{x}$  as

$$\tilde{\mathbf{X}} = \begin{bmatrix} x_0 & x_1 & \cdots & x_{J-2} & x_{J-1} \\ x_1 & x_2 & \cdots & x_{J-1} & x_0 \\ x_2 & x_3 & \cdots & x_0 & x_1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{2N-1} & x_{2N} & \cdots & x_{2N-3} & x_{2N-2} \end{bmatrix}_{2N \times J} \quad (\text{A.45})$$

and we consider the power (energy divided by the number of coefficients) of the low-pass and high-pass filter outputs. Since the number of coefficients is  $J/2$ , due to the down sampling operation, the power values of the approximation ( $\mathbf{c}^m$ ) and of the detail ( $\mathbf{d}^m$ ) coefficients for the  $m^{\text{th}}$  training signal are given by

$$P_c^m = \frac{\mathbf{c}^m \mathbf{c}^{mT}}{J/2} \quad (\text{A.46a})$$

$$P_d^m = \frac{\mathbf{d}^m \mathbf{d}^{mT}}{J/2} \quad (\text{A.46b})$$

If  $M$  training signals are employed, the overall power of the approximations and details are

$$P_c = \sum_{i=1}^M P_c^m \quad (\text{A.47a})$$

$$P_d = \sum_{i=1}^M P_d^m \quad (\text{A.47b})$$

The relation between  $P_c$  and  $P_d$  can be expressed in an objective function  $F: \mathfrak{R}^{2N} \times \mathfrak{R}^{2N} \rightarrow \mathfrak{R}$  given by:

$$F(\mathbf{h}, \mathbf{g}) = \frac{0.5(P_c + P_d)}{\sqrt{P_c P_d}} \quad (\text{A.48})$$

which is similar to the coding gain used to assess the compression performance of two-channel filter bank structures [31].

If the conditions in Eq.(A.34) are satisfied, the WT is orthogonal, and thus it preserves power [33]. As a result, the power of the input signal  $\mathbf{x}^m$  equals the sum of the power in both output channels,  $P_c^m + P_d^m$ . Hence, for a given training set, the sum  $A = P_c + P_d$  is constant and thus maximising  $F$  is equivalent to maximising  $P_c$ . Writing:

$$F^2 = \frac{0.25A^2}{AP_c - P_c^2} \quad (\text{A.49})$$

it follows that  $F$  reaches a minimum of 1 when  $P_c = A/2$ , that is, when the power is equally divided between the approximation and detail coefficients. As  $P_c$  increases from  $A/2$  to  $A$ ,  $F$  increases and tends to  $+\infty$  when  $P_c$  tends to  $A$ , i.e, when all the power is contained in the approximation coefficients. The advantage of aiming at maximising  $P_c$  lies in the fact that the signal-to-noise ratio is usually larger in the low-pass filter output. Thus, maximising  $P_c$  further improves the filtering performance. It is worth noting that, if the data set is mean-centered (variable-wise) prior to the wavelet decomposition, the power is equal to the variance. In this manner, the optimisation amounts to maximizing the variance explained by the approximation coefficients.

Since  $P_c$  only depends on the low-pass filter weights  $\mathbf{h}$ , the problem can be re-stated as the maximization of an objective function  $\varepsilon: \Re^{2N} \rightarrow \Re$  given by:

$$\varepsilon(\mathbf{h}) = P_c = \frac{1}{J/2} \sum_{m=1}^M \mathbf{c}^m \mathbf{c}^{mT} \cong \frac{1}{J} \sum_{i=1}^M \mathbf{c}^m \mathbf{c}^{mT} \quad (\text{A.50})$$

where  $\mathbf{c}^m = \mathbf{h} \tilde{\mathbf{X}}^m$  is the vector of detail coefficients for the  $m^{\text{th}}$  training signal, before the downsampling operation as shown in Table A.1 and  $\tilde{\mathbf{X}}^m$  is the circulant matrix formed from  $\mathbf{x}^m$ . This holds because we may assume that the power of the detail coefficients before and after downsampling is approximately the same. Using (A.50) one may write:

$$\varepsilon(\mathbf{h}) = \frac{1}{J} \sum_{m=1}^M \mathbf{h} \tilde{\mathbf{X}}^m \tilde{\mathbf{X}}^{mT} \mathbf{h}^T = \mathbf{h} \underbrace{\sum_{m=1}^M \left( \frac{1}{J} \tilde{\mathbf{X}}^m \tilde{\mathbf{X}}^{mT} \right)}_{\mathbf{R}} \mathbf{h}^T \quad (\text{A.51})$$

Since  $\tilde{\mathbf{X}}^m \tilde{\mathbf{X}}^{mT}$  is Toeplitz for any  $\mathbf{x}^m$ , then  $\mathbf{R}_{2N \times 2N}$  is also a Toeplitz matrix. Thus, the constraints in Equation (A.34) allow the objective function to be rewritten, with a slight abuse of notation, in the following linear form [34]

$$\varepsilon(\mathbf{a}) = \frac{r_0}{2} + \sum_{n=0}^{N-1} a_n r_{2n+1} \quad (\text{A.52})$$

where  $\{r_0, r_1, \dots, r_{2N-1}\}$  are the elements of the first row of  $\mathbf{R}$  and vector  $\mathbf{a} = [a_0 \ a_1 \ \dots \ a_{N-1}]$  contains the coefficients of the product filter  $P(z)$  defined as:

$$P(z) = H(z)H(z^{-1}) = 1 + \sum_{n=0}^{N-1} a_n (z^{2n-1} + z^{2n+1}) \quad (\text{A.53})$$

Given  $\mathbf{a}$ , the transfer function  $H(z)$  of the desired filter can be recovered from  $P(z)$  by a spectral factorisation procedure [33]. This factorisation is possible provided that the frequency response of the product filter given by  $Q(f) = P(e^{j2\pi f})$  (where  $j$  is the imaginary unity), is non-negative at all frequencies  $f$ , that is,  $Q(f) \geq 0$ ,  $\forall f \in \mathfrak{R}$ . It follows that the following restriction must be enforced:

$$Q(f) = 1 + 2 \sum_{n=0}^{N-1} a_n \cos[2\pi f(2n+1)] \geq 0 \quad (\text{A.54})$$

Since  $Q(f)$  is periodic with period 1 and  $Q(f) + Q(f+0.5) = 2$ ,  $\forall f \in \mathfrak{R}$ , it is sufficient to consider the restriction  $Q(f) \geq 0$  in the interval  $0 \leq f \leq 0.5$ , that is

$$1 + 2 \sum_{n=0}^{N-1} a_n \cos[2\pi f(2n+1)] \geq 0, \quad 0 \leq f \leq 0.5 \quad (\text{A.55})$$

Maximising  $\varepsilon(\mathbf{a})$  defined in Equation (A.52) with respect to  $\mathbf{a}$  subject to the inequality restrictions in Equation (A.55) is a linear semi-infinite programming (LSIP) problem [35], because there is a finite number of variables ( $a_0 \ a_1 \ \dots \ a_{N-1}$ ) and infinitely many restrictions. This problem can be solved by discretising the frequency interval  $[0, 0.5]$  to generate a finite number of restrictions, and then applying standard linear programming techniques [36]. The solution  $\tilde{\mathbf{a}}$  to this approximated problem can then be used to generate a feasible solution  $\mathbf{a}_f$  to the original problem as discussed by Eriksson *et al.*, (2000) [34]:

$$\mathbf{a}_f = \frac{\tilde{\mathbf{a}}}{1 - \delta} \quad (\text{A.56})$$

where  $\delta \leq 0$  is the minimum of  $Q(f)$  in the interval  $0 \leq f \leq 0.5$  when  $\tilde{\mathbf{a}}$  is used instead of  $\mathbf{a}$  in Equation (A.54).

From the above description, it may be concluded that by adopting the coding gain as a measure of the compression performance of the low-pass/high-pass filter pair [37], the optimisation of coefficients  $\{a_0, a_1, \dots, a_{N-1}\}$  can be cast into a LSIP problem. By using a convenient discretisation procedure [32] such a problem can then be converted into a linear programming one, for which efficient solution algorithms exist.

#### A.4.5 Parametrisation and optimisation approach

The method of parameterisation of perfect reconstruction finite impulse response filter banks and adapts it to parameterise orthonormal wavelets of arbitrary compact support have been stated by Vaidyanathan's [29]. Moreover, the literature discusses several parameterisations of orthonormal wavelets. In [38] Zou and Tewfik (1993) proposed a parameterise wavelets method that can be used for parameterise wavelets with a number of vanishing moments greater than one.

Let:

$$\begin{aligned} h_0^{(1)} &= \cos(a_1), & h_1^{(1)} &= \sin(a_1) \\ h_0^{(N+1)} &= \cos(a_{N+1})h_0^{(N)} & h_1^{(N+1)} &= \sin(a_{N+1})h_0^{(N)} \\ h_{2i}^{(N+1)} &= \cos(a_{N+1})h_{2i}^{(N)} - \cos(a_{N+1})h_{2i-1}^{(N)}, \quad i=1, 2, \dots, N-1 & h_{2i+1}^{(N+1)} &= \sin(a_{N+1})h_{2i}^{(N)} + \cos(a_{N+1})h_{2i-1}^{(N)}, \quad i=1, 2, \dots, N-1 \\ h_{2N}^{(N+1)} &= -\sin(a_{N+1})h_{2N-1}^{(N)} & h_{2N+1}^{(N+1)} &= \cos(a_{N+1})h_{2N-1}^{(N)} \end{aligned}$$

As stated in [39] [40], in order to ensure two vanishing moments for the resulting transform,

$$a_N = \frac{\pi}{4} - \sum_{i=1}^{N-1} a_i \quad (\text{A.57})$$

$$a_{N-1} = \frac{1}{2} \arcsin \left\{ -\frac{1}{2} - \sum_{k=1}^{N-2} \left[ \sin \sum_{i=1}^k 2a_i \right] \right\} - \sum_{i=1}^{N-2} a_i \quad (\text{A.58})$$

In what follows, an optimization process is proposed that maximizes the selectivity of the pair of high-pass/low-pass orthonormal wavelet filters with a given length.

As discussed in [39],[41] the expression in (A.57) has a real value solution if the set of angles  $a_i$  where  $1 \leq i \leq N-2$  satisfy a set of constrains that define a non-convex region in  $R^{N-2}$ . Additional constrains are imposed to ensure this, by invoking a new parameter  $\chi_i$  so that:

$$\chi_i = \sin \sum_{k=1}^i 2a_k, \quad 1 \leq i \leq N-2 \quad (\text{A.59})$$

$$-\frac{3}{2} \leq \sum_{i=1}^{N-2} \chi_i \leq \frac{1}{2}, \quad -1 \leq \chi_1, \chi_2 \cdots \chi_{N-2} \leq 1 \quad (\text{A.60})$$

Rearranging (A.58) and (A.59) we have:

$$a_1 = \frac{1}{2} \arcsin(\chi_1) \quad (\text{A.61})$$

$$a_i = \frac{1}{2} \arcsin(\chi_i) - \sum_{k=1}^{i-1} 2a_k, \quad 2 \leq i \leq N-2 \quad (\text{A.62})$$

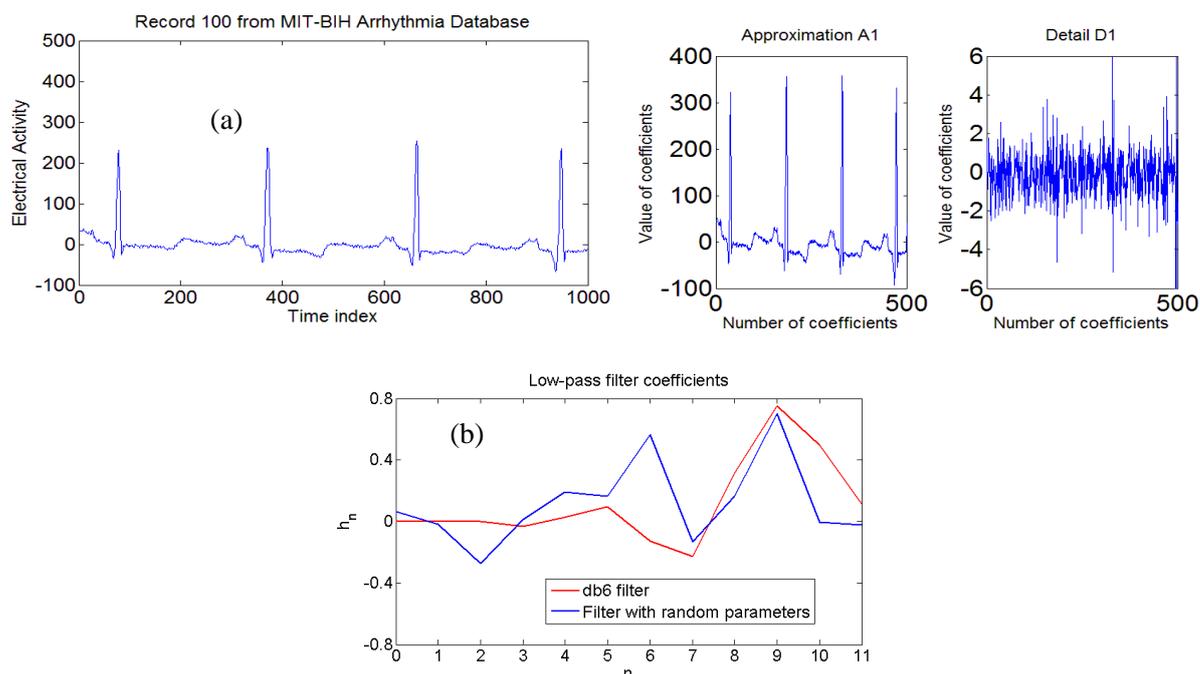
The cost  $J$  is defined from the frequency response of the low pass filter as:

$$J = \frac{\int_{0.25\omega_s}^{0.5\omega_s} |H^{(N)} e^{j\omega T}| d\omega}{\int_0^{0.5\omega_s} |H^{(N)} e^{j\omega T}| d\omega} \quad (\text{A.63})$$

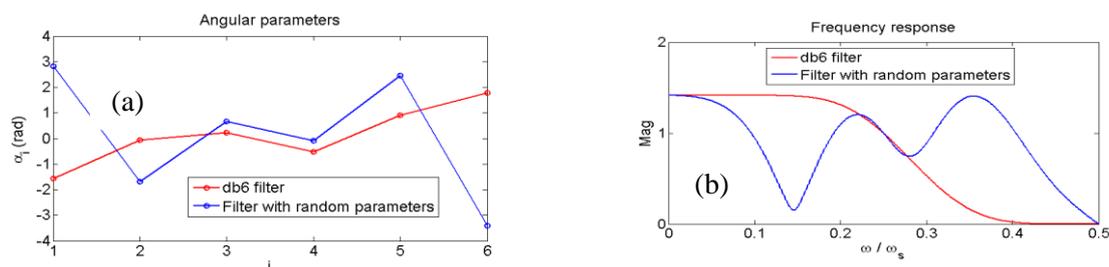
with  $\omega_s = 2\pi/T$  being the sampling frequency. As discussed in [41], optimization of the cost function with respect to the  $\chi_i$  parameters is accomplished by using sequential quadratic programming (SQP) which uses local quadratic approximations of the cost function and local linear approximations of the restrictions. An interesting variance to the above algorithm where additional constrains are imposed to ensure a third vanishing moment can be found in [41].

#### A.4.6 Example of a signal decomposition process

The signal in Fig. A.16 represents the first 1000 points out of a 3600 data points record from patient number 100 (lead 1) from the MIT database. The patient does not have a pathogenic condition and his record (among others from that database) is normally used as a training set to different classifiers to discriminate from other pathogenic patient records. A typical decomposition of the signal to approximation and detail coefficients at the first decomposition level is shown as an inset to that figure. Normally, a much smaller filter tap is generated by the user as shown in the Figure A.16b. Figure A.17 depicts the angular parameter alpha associated with a standard db6 filter bank as well as for the filter bank generated on the basis of the proposed procedure. Figure A.17b shows the difference in the function of these filters in the frequency domain. The introduction of a vanishing moment in the random parameter filter shown in Fig. A.17b ensures that its gain drops to zero at high frequencies.



**Figure A.16:** a) Typical signal from the MIT-BIH database with corresponding reconstruction on the basis of approximation and wavelet coefficients at the first decomposition level and b) comparison of filter coefficients impulse response function assuming 12 taps.



**Figure A.17:** a) Comparison of angular parameters for a standard db6 filter bank with those of a filter bank generated using the proposed procedure and b) normalized frequency response for the two filter banks depicting the difference in their function.

## A.5 Hardware requirements in ECG processing

In (2012) Rai and Trivedi indicated that ECG signal-processing techniques consist of de-noising baseline correction, parameter extraction and arrhythmia detection [42]. The extraction of a heartbeat characteristic waves in a time resolved manner are the main requirements for all ECG signal-processing algorithm. Execution of the digital signal processing using specifically designed hardware for the purpose of assisting physicians or cardiologists performing ECG monitoring has been discussed in [43]. A real-time visual observation of the signal on LCD screen is provided in many wearable devices nowadays.

Three hardware units are needed in the design of any the standalone DSP hardware. These are: (1) a digital signal processor with memory unit; (2) a signal monitoring displays unit; and (3) an analogue front-end unit. These units are integrated through firmware. An ideal system should have the ability to perform in standalone mode and be upgradeable to account for advancement in algorithmic development. Additionally, it should be simple to operate and have low power consumption, and should not be expensive. The hardware should provide as well data storage and a communication protocol. Its main components include: electrical erasable programmable read-only memory (EEPROM – CAT28LV64W), a floating point DSP processor (TMS320VC33), and static random access memory (SRAM – CY7C1041DV33). In the signal monitoring display unit, the input peripherals are used for the observation of the input signal. The PIC18F452 microcontroller can be set up to communicate with graphical liquid crystal displays (GLCD). The signal conditioning circuit (SCC) can processes the ECG signal by converting from its analog to its digital form using an ADC0820 chip. The output port (DB0 to DB7) of the ADC0820 chip a provide the converted ADC data. The analogue front-end unit can provide communication with the real-world by means of the peripherals connectivity. In such system, the digital signal processor may be used as the core processing unit in order to establish standalone DSP hardware. The graphic LCD normally is used for the visual observation of ECG signals. The algorithm for this system can still be changed or updated so as to provide improvement in performance [44].

## **A.6 Summary**

This appendix reviewed the continuous and discrete wavelet transforme and associated filter banks for ECG signal analysis. Filters with vanishing moments of higher order were discussed. Filter banks such as biorthogonal and orthogonal are considered as the arrangements of bandpass, high pass and lowpass filters can be used to divide the data into subbands. The wavelet transform has the capability to localize events in time and frequency. and can perform in non-stationary signal analysis. wavelet transforms can produce a finite energy function along an orthonormal basis. An ECG test waveform was decomposed to qualitatively describe the new parametrization process being implemented. This decomposition was contrasted to that of a standard db6 parametrization. Further work intends to optimize the number of wavelet coefficients as well as number of decomposition levels presented to the classifiers discussed in subsequent chapter 6.

## A.7 References related to the wavelet discussion in the Appendix

- [1] O. Ondráček, J. Púčik, and E. Cocherová, “Filters for Ecg Digital Signal Processing,” *International Conference ” Trends in Biomedical Engineering” September 7 -9 ,2005, University of Zilina*, pp. 91–96, 2005.
- [2] C.L. Liu, “A Tutorial of the Wavelet Transform,” NTUEE, Taiwan, 2010.
- [3] L. A. Conraria and M. J. Soares, “The Continuous Wavelet Transform: A Primer,” *NIPE Working Paper*, vol. 16, pp. 1–43, 2011.
- [4] R. Panda, “Removal of Artifacts from Electrocardiogram,” National Institute of Technology Rourkela, 2012.
- [5] J.P. Antoine, P. Carrette, R. Murenzi, and B. Piette, “Image analysis with two-dimensional continuous wavelet transform,” *Signal Processing*, vol. 31, no. 3, pp. 241–272, 1993.
- [6] H. Zheng, Z. Li, and X. Chen, “Gear Fault Diagnosis Based on Continuous Wavelet Transform,” *Mechanical Systems and Signal Processing*, vol. 16, no. 2–3, pp. 447–457, 2002.
- [7] P. Du, W. a. Kibbe, and S. M. Lin, “Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching,” *Bioinformatics*, vol. 22, no. 17, pp. 2059–2065, 2006.
- [8] S. Yunhui and R. Qiuqi, “Continuous wavelet transforms,” in *Signal Processing, 2004. Proceedings. ICSP ’04. 2004 7th International Conference on*, 2004, vol. 1, pp. 207–210.
- [9] A. Bultheel and D. Huybrechs, “Wavelets with applications in signal and image processing,” *Course material University of Leuven, Belgium*, 2011.
- [10] M. Kociolek, A. Materka, M. Strzelecki, and P. Szczypinski, “Discrete Wavelet Transform – Derived Features for Digital Image Texture Analysis,” in *International Conference on Signals and Electronic Systems*, 2001, pp. 163–168.
- [11] J. Sadowsky, “Investigation of Signal Characteristics Using the Continuous Wavelet Transform,” *John Hopkins APL Technical Digest*, vol. 17, no. 3, pp. 258–269, 1996.
- [12] C. . S. Burrus, R. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transformer: A Primer*. United States of America: Prentice Hall, 1998.
- [13] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelet Toolbox™ User ’ s Guide*. The MathWorks, Inc., 2016.
- [14] Z. Hong-tu and Y. Jing, “The Wavelet Decomposition And Reconstruction Based on The Matlab,” in *Proceedings of the Third International Symposium on Electronic Commerce and Security Workshops (ISECS ’10)*, 2010, pp. 143–145.
- [15] Z. Hong-tu and Y. Jing, “The Wavelet Decomposition And Reconstruction Based on The Matlab,” in *Proceedings of the Third International Symposium on Electronic Commerce and Security Workshops (ISECS ’10)*, 2010, pp. 143–145.
- [16] J. Rafiee, M. a. Rafiee, and P. W. Tse, “Application of mother wavelet functions for automatic gear and bearing fault diagnosis,” *Expert Systems with Applications*, vol. 37, no. 6, pp. 4568–

- 4579, 2010.
- [17] R. A. Gopinath and C. S. Burrus, "Wavelet transforms and filter banks," in *Wavelets: a tutorial in theory and applications*, C. K. Chui, Ed. Academic Press Professional, Inc., 1992, pp. 603–654.
- [18] C. Herley, "Wavelets and Filter banks," Columbia University, 1995.
- [19] S. T. Prasad and S. Varadarajan, "ECG Signal Processing Using Digital Signal Processing Techniques," *International Journal of Scientific & Engineering Research*, vol. 4, no. 12, pp. 1624–1628, 2013.
- [20] D. Balasubramaniam and D. Nedumaran, "Implementation of ECG signal processing and analysis techniques in digital signal processor based system," *IEEE International Workshop on Medical Measurements and Applications*, pp. 60–63, 2009.
- [21] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial," *Proceedings of the IEEE*, vol. 78, no. 1, pp. 56–93, 1990.
- [22] R. S. S. Kumari, S. Bharathi, and V. Sadasivam, "Design of Optimal Discrete Wavelet for ECG Signal Using Orthogonal Filter Bank," in *International Conference on Computational Intelligence and Multimedia Applications 2007 (ICCIMA 2007)*, 2007, vol. 1, pp. 525–529.
- [23] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "ECG beat detection using filter banks," *IEEE Transactions on Biomedical Engineering*, vol. 46, no. 2, pp. 192–202, 1999.
- [24] V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo, "Multirate processing of the ECG using filter banks," *Computers in Cardiology*, pp. 245–248, 1996.
- [25] M. M. Sadaphule, P. S. B. Mule, and P. S. O. Rajankar, "ECG Analysis Using Wavelet Transform and Neural Network," *International Journal of Engineering Inventions*, vol. 1, no. 12, pp. 1–7, 2012.
- [26] T. Froese, S. Hadjiloucas, R. K. H. Galvão, V. M. Becerra, and C. J. Coelho, "Comparison of extrasystolic ECG signal classifiers using discrete wavelet transforms," *Pattern Recognition Letters*, vol. 27, no. 5, pp. 393–407, 2006.
- [27] N. Jannah, S. Hadjiloucas, F. Hwang, and R. K. H. Galvão, "Smart-phone based electrocardiogram wavelet decomposition and neural network classification," *Journal of Physics: Conference Series*, vol. 450, pp. 1–7, 2013.
- [28] B. G. Sherlock and D. M. Monro, "On the space of orthonormal wavelets," *IEEE Transactions on Signal Processing*, vol. 46, no. 6, pp. 1716–1720, 1998.
- [29] P. Vaidyanathan, *Multirate Systems and Filter Banks*. New Jersey, United States of America: Prentice-Hall, Inc, 1993.
- [30] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. United States of America: Wellesley-Cambridge Press Massachusetts., 1996.
- [31] J. Tuqan and P. Vaidyanathan, "A state space approach to the design of globally optimal FIR energy compaction filters," *IEEE transactions on signal processing*, vol. 48, no. 10, pp. 2822–2838, 2000.

- [32] P. Moulin, M. Anitescu, K. O. Kortanek, and F. A. Potra, "Role of linear semi-infinite programming in signal-adapted QMF bank design," *IEEE Transactions on Signal Processing*, vol. 45, no. 9, pp. 2160–2174, 1997.
- [33] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley-Cambridge: Wellesley-Cambridge Press Massachusetts., 1996.
- [34] L. Eriksson, J. Trygg, E. Johansson, R. Bro, and S. Wold, "Orthogonal signal correction, wavelet analysis, and multivariate calibration of complicated process fluorescence data," *Analytica Chimica Acta*, vol. 420, no. 2, pp. 181–195, 2000.
- [35] R. Hettich and K. O. Kortanek, "Semi-infinite programming - theory, methods, and applications," *SIAM Review*, vol. 35, no. 3, pp. 380–429, 1993.
- [36] V. Chvatal, *Linear Programming*. New York: W.H.Freeman and Company, 1983.
- [37] M. Unser, "On the Optimality of Ideal Filters for Pyramid and Wavelet Signal Approximation," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3591–3596, 1993.
- [38] H. Zou and A. H. Tewfik, "Parametrization of Compactly Supported Orthonormal Wavelets," *IEEE Transactions on Signal Processing*, vol. 41, no. 3, pp. 1428–1431, 1993.
- [39] H. M. Paiva, M. N. Martins, R. K. H. Galvão, and J. P. L. M. Paiva, "On the space of orthonormal wavelets: Additional constraints to ensure two vanishing moments," *IEEE Signal Processing Letters*, vol. 16, no. 2, pp. 101–104, 2009.
- [40] H. M. Paiva and R. K. H. Galvão, "Optimized orthonormal wavelet filters with improved frequency separation," *Digital Signal Processing*, vol. 22, no. 4, pp. 622–627, 2012.
- [41] J. C. Uzinski, H. M. Paiva, F. Villarreal, M. A. Q. Duarte, and R. K. H. Galvão, "Additional constraints to ensure three vanishing moments for orthonormal wavelet filter banks," *Anais do Congresso de Matemática Aplicada e Computacional*, pp. 16–19, 2013.
- [42] H. M. Rai and A. Trivedi, "Classification of ECG Waveforms for Abnormalities Detection using DWT and Back Propagation Algorithm," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 1, no. 4, pp. 517–520, 2012.
- [43] G. C. Seng, S. Salleh, A. R. Harris, J. M. Najeb, and I. Kamarulafizam, "Standalone ECG Monitoring System Using Digital Signal Processing Hardware," in *Progress In Electromagnetics Research Symposium Proceedings, KL, Malaysia, March 27-30, 2012*, pp. 1901–1904.

## Appendix B Publications presented during the doctoral research

The following journal, publications and conference papers that have been presented during the doctoral research and part of the work in thesis has been taken from these publications:

Jannah, N., Hadjiloucas, S., Hwang, F., & Galvão, R. K. H. (2013). Smart-phone based electrocardiogram wavelet decomposition and neural network classification. *Journal of Physics: Conference Series*, 450, 012019.

Hadjiloucas, S., Jannah, N., Hwang, F., & Galvão, R. K. H. (2014). On the application of optimal wavelet filter banks for ECG signal classification. *Journal of Physics: Conference Series*, 490, 012142.

Jannah, N., & Hadjiloucas, S. (2015). Support Vector Machine Classification of ECG Signal with Pre-processing in the Principle Component and Wavelet Domains . In *8th Saudi Students Conference in the UK - 2015 Abstracts of Accepted Posters in Engineering and Applied Sciences Posters*

<http://uksacb.org/ssc8/conference/programme/50-posters/174-engineering-and-applied-sciences-posters>

Jannah, N., & Hadjiloucas, S. (2015). Detection of ECG Arrhythmia Conditions using CSVM and MSVM Classifiers. In *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1–2 Philadelphia: IEEE.

Jannah, N., & Hadjiloucas, S. (2016). ECG Arrhythmia Classifiers : Comparison between Multi-class SVM and CSVM Algorithms., *9th Saudi Students Conference in the UK - 2016 Abstracts of Accepted Papers and Posters* (p. 481). The Scientific Society for Saudi Students in the UK, 2016. Retrieved from [http://uksacb.org/sites/default/files/user/SSC9-Proceedings\\_0.pdf](http://uksacb.org/sites/default/files/user/SSC9-Proceedings_0.pdf)

Jannah, N., & Hadjiloucas, S. (2016). A novel technique for classifier cardiac arrhythmia condition using SIMCA. In *Sensors & their Applications XVIII 2016, 12-13 September 2016*

Jannah, N., & Hadjiloucas, S. (2017). A Comparison Between ECG Beat Classifiers Using Multiclass SVM and SIMCA with Time Domain PCA Feature Reduction. In *UKSim-AMSS 19th International Conference on Mathematical Modelling & Computer Simulation*, pp. 126–131.

### Project related award

Poster paper prize from IoP (Institute of Physics) ISAT Group (Instrument Science and Technology) - Najlaa Jannah- for the poster presentation:

Jannah, N., & Hadjiloucas, S. (2016). A novel technique for classifier cardiac arrhythmia condition using SIMCA. In *Sensors & their Applications XVIII 2016, 12-13 September 2016*

**IOPscience** iopscience.iop.org

[Home](#) [Search](#) [Collections](#) [Journals](#) [About](#) [Contact us](#) [My IOPscience](#)

## Smart-phone based electrocardiogram wavelet decomposition and neural network classification

This content has been downloaded from IOPscience. Please scroll down to see the full text.

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

### Download details:

IP Address: 134.225.206.51

This content was downloaded on 04/10/2016 at 11:04

Please note that [terms and conditions apply](#).

You may also be interested in:

[Robust electrocardiogram \(ECG\) beat classification using discrete wavelet transform](#)

Fayyaz-ul-Amir Afsar Minhas and Muhammad Arif

[Application of the phasor transform for automatic delineation of single-lead ECG fiducial points](#)

Arturo Martínez, Raúl Alcaraz and José Joaquín Rieta

[System identification algorithms for the analysis of dielectric responses from broadband spectroscopies](#)

S Hadjiloucas, G C Walker, J W Bowen et al.

[Real-time detection of pathological cardiac events in the electrocardiogram](#)

Ivo Iliev, Vessela Krasteva and Serafim Tabakov

[Wavelet bidomain sample entropy in AF](#)

Raúl Alcaraz and José Joaquín Rieta

[Development of an algorithm for heartbeats detection and classification in Holter records based on temporal and morphological features](#)

A García, H Romano, E Laciari et al.

**IOPscience** iopscience.iop.org

[Home](#) [Search](#) [Collections](#) [Journals](#) [About](#) [Contact us](#) [My IOPscience](#)

## On the application of optimal wavelet filter banks for ECG signal classification

This content has been downloaded from IOPscience. Please scroll down to see the full text.

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

### Download details:

IP Address: 134.225.206.51

This content was downloaded on 04/10/2016 at 11:03

Please note that [terms and conditions apply](#).

You may also be interested in:

[Smart-phone based electrocardiogram wavelet decomposition and neural network classification](#)

N Jannah, S Hadjiloucas, F Hwang et al.

[Wavelet based detection of changes in the composition of RLC networks](#)

H M Paiva, M A Q Duarte, R K H Galvao et al.

[III Lead ECG Pulse Measurement Sensor](#)

S K Thangaraju and K Munisamy

[ECG feature extraction based on the bandwidth properties of variational mode decomposition](#)

Ahmet Mert

[Matrix of regularity for improving the quality of ECGs](#)

Henian Xia, Gabriel A Garcia, Jujhar Bains et al.

[Phase retrieval from a single fringe pattern by using empirical wavelet transform](#)

Xiaopeng Guo, Hong Zhao and Xin Wang

[De-noising of two-dimensional angular correlation of positron annihilation radiation data using](#)

[Daubechies wavelet thresholding](#)

A G Major, H M Fretwell, S B Dugdale et al.

# On the application of optimal wavelet filter banks for ECG signal classification

S. Hadjiloucas<sup>1</sup> | N. Jannah<sup>1</sup> | R.K.H. Galvão<sup>2</sup> |

## Introduction

This paper discusses ECG classification after parametrizing the ECG waveforms in the wavelet domain. Signal decomposition using perfect reconstruction quadrature mirror filter banks can provide a very parsimonious representation of ECG signals. In the current work, the filter parameters are adjusted by a numerical optimization algorithm in order to minimize a cost function associated to the filter cut-off sharpness. The goal consists of achieving a better compromise between frequency selectivity and time resolution at each decomposition level than standard orthogonal filter banks such as those of the Daubechies and Coiflet families.

## Wavelet filter bank parametrization.

In this filter bank, the low-pass filtering result undergoes successive filtering iterations with the number of iterations  $N_{it}$  chosen by the analyst. The final result of the decomposition of data vector  $\mathbf{x}$  is a vector resulting from the concatenation of row vectors  $\mathbf{c}(N_{it})$  (termed approximation coefficient at the largest scale level) and  $\mathbf{d}(s)$  (termed detail coefficients at the  $s^{\text{th}}$  scale level,  $s = 1, \dots, N_{it}$ ) in the following manner:

$$\mathbf{t} = [\mathbf{c}(N_{it}) | \mathbf{d}(N_{it}) | \mathbf{d}(N_{it}-1) | \dots | \mathbf{d}(1)]$$

The filter bank transform can be regarded as a change in variables from  $\mathfrak{R}$  to  $\mathfrak{R}'$  performed according to the following operation,

$$t_j = \sum_{n=0}^{J-1} x_n v_j(n), \quad j = 0, 1, \dots, J-1$$

Where  $t_j$  is a transformed variable and  $v_j(n) \in \mathfrak{R}$  is a transform weight. It proves convenient to write the transform in matrix form as:

$$\mathbf{t}_{1 \times J} = \mathbf{x}_{1 \times J} \mathbf{V}_{J \times J}$$

where  $\mathbf{x} = [x_0 \ x_1 \ \dots \ x_{J-1}]$  is the row vector of original variables,  $\mathbf{t}$  is the row vector of new (transformed) variables and  $\mathbf{V}$  is the matrix of weights.

$$\mathbf{V} = \begin{bmatrix} 0 & 0 & \dots & h_{2N-4} & h_{2N-2} & 0 & 0 & \dots & \varepsilon_{2N-4} & \varepsilon_{2N-2} \\ h_{2N-1} & 0 & \dots & h_{2N-3} & h_{2N-1} & \varepsilon_{2N-4} & 0 & \dots & \varepsilon_{2N-3} & \varepsilon_{2N-1} \\ h_{2N-2} & 0 & \dots & h_{2N-4} & h_{2N-2} & \varepsilon_{2N-2} & 0 & \dots & \varepsilon_{2N-1} & \varepsilon_{2N-4} \\ \vdots & \vdots \\ h_0 & h_0 & \dots & 0 & 0 & \varepsilon_0 & \varepsilon_1 & \dots & 0 & 0 \\ 0 & h_0 & \dots & 0 & 0 & 0 & \varepsilon_0 & \dots & 0 & 0 \\ 0 & h_0 & \dots & 0 & 0 & 0 & \varepsilon_0 & \dots & 0 & 0 \\ \vdots & \vdots \\ 0 & 0 & \dots & h_{2N-2} & 0 & 0 & 0 & \dots & \varepsilon_{2N-2} & 0 \\ 0 & 0 & \dots & h_{2N-3} & h_{2N-1} & 0 & 0 & \dots & \varepsilon_{2N-3} & \varepsilon_{2N-1} \end{bmatrix}$$

Filtering is carried out by circular convolution[1].

$x_0$	$x_1$	...	$x_{2N-1}$	$x_{2N}$	...	$x_{J-1}$	$x_0$	$x_1$	...	$x_{2N-2}$	Before	After
$h_{2N-1}$	$h_{2N-2}$	...	$h_0$								$c_0'$	$c_0$
	$h_{2N-1}$	...	$h_1$	$h_0$							$c_1'$	$c_1$
		...	$\vdots$								$\vdots$	$\vdots$
							$h_{2N-1}$	$h_{2N-2}$	$h_{2N-3}$	...	$c_{J-2}'$	
							$h_{2N-1}$	$h_{2N-2}$	...	$h_0$	$c_{J-1}'$	$c_{J-1}$

Table 1. Circular convolution procedure for low-pass filtering showing results before and after dyadic down-sampling. The detail coefficients  $d_s$  are obtained in a similar manner by using the high-pass filtering sequence.

## Parametrization and optimization approach.

An optimization process that maximizes the selectivity of the pair of high-pass/low-pass orthonormal wavelet filters with a given length. The following expression has a real value solution if the set of angles  $a_i$  where  $1 \leq i \leq N-2$  satisfy a set of constrains that define a non-convex region in  $R^{N-2}$ .

$$a_{N-1} = \frac{1}{2} \arcsin \left[ -\frac{1}{2} - \sum_{k=1}^{N-2} \left[ \sin \sum_{i=1}^k 2a_i \right] \right] - \sum_{i=1}^{N-2} a_i$$

Additional constrains are imposed to ensure this, by invoking a new parameter so that:

$$x_i = \sin \sum_{k=1}^{N-2} 2a_k, \quad 1 \leq i \leq N-2 \quad -\frac{3}{2} \leq \sum_{i=1}^{N-2} x_i \leq \frac{1}{2}, \quad -1 \leq x_1, x_2, \dots, x_{N-2} \leq 1$$

The cost  $J$  is defined from the frequency response of the low pass filter as following expression [2,3]:

$$J = \frac{\int_0^{0.5\omega_s} |H(N)e^{j\omega T}| d\omega}{\int_0^{0.5\omega_s} |H(N)e^{j\omega T}| d\omega}$$

Our aim is to optimally decompose the signals in the wavelet domain (function  $J$ ) so that they can be subsequently used as inputs for training to a recurrent neural network classifier.

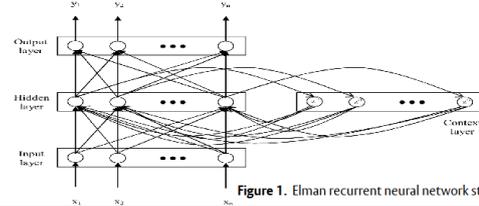


Figure 1. Elman recurrent neural network structure.

### References

1. Froese, T, Hadjiloucas, S, Galvão, R K H, Becerra V M and Coelho, C J, 'Comparison of extrasystolic ECG signal classifiers using Discrete Wavelet Transforms' *Pattern Recognition Letters* 27(5) 393-407.
2. Paiva H M, Marins M N, Galvão R K H and Paiva J P L M, 'On the space of orthonormal wavelets: additional constrains to ensure two vanishing moments', *IEEE Signal Processing Letters*, 16, 101-104.
3. Paiva H M and Galvão R K H, 'Optimized orthonormal wavelet filters with improved frequency separation', *Digital Signal Processing*, 22, 622-627.

### Contact information

- <sup>1</sup>School of Systems Engineering, University of Reading, Whiteknights, RG6 6AH, UK.
- <sup>2</sup>Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, Brazil.
- Work supported by CNPq (research fellowships) and FAPESP (grants 2011/13777-8, 2011/17610-0) and a Saudi Arabia scholarship
- Email: s.hadjiloucas@reading.ac.uk | Sillas Hadjiloucas



# Support Vector Machine Classification of ECG Signal with Pre-processing in the Principle Component and Wavelet Domains

N. Jannah | S. Hadjloucas |

## Introduction

Cardio-vascular disease is one of the leading causes of death worldwide. Early diagnoses of heart disease can save the lives of many people, and this can be accomplished through the analysis of a electrocardiogram (ECG). ECG analysis is one of the most well established diagnosis tools for heart condition and arrhythmia detections. The ECG signal represents the changes in electrical potential during the cardiac cycle. These are recorded using surface electrodes on the body (arm, leg, and chest wall). This work evaluates ECG classification using pre-processing routines, feature extraction and selection, and ECG beat classification using Support Vector Machine (SVM). The basic components of this study consist of ECG beat detection, data pre-processing for feature selection by using wavelet coefficients from discrete wavelet transform and PC coefficients from Principal Component Analysis (PCA), and also beat classification using both binary and multi-class SVM.

## Method and Procedure

Diagrams below summarises the method required in each step of performing ECG classification.

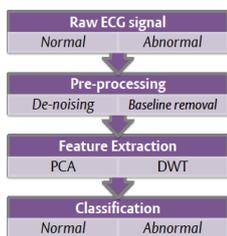


Figure (1) Structure and methodology of ECG classification.

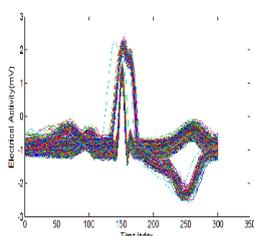


Figure (2) Extract of beats from one lead

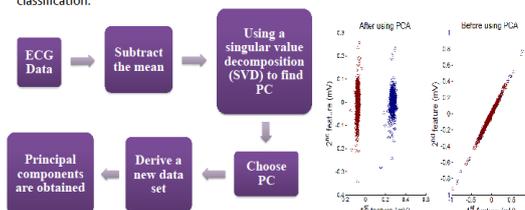


Figure (3) Procedure of PCA

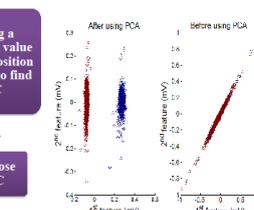


Figure (4) Compare between feature before and after using PCA

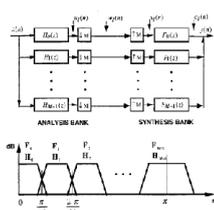


Figure (5) Analysis and synthesis filter banks [1]

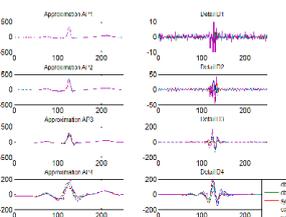


Figure (6) Normal ECG signal multilevel decomposition using some wavelet families.

## Implementation using LIBSVM and Results

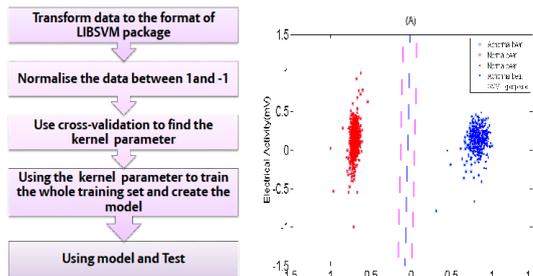


Figure (7) Procedure of LIBSVM classifiers

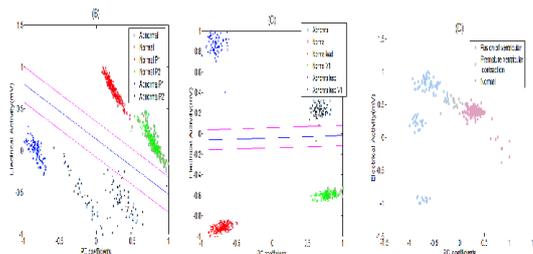


Figure (8) Classification result using SVM and ECG signal that taken from: (A) one lead (B) two patients (C) two leads and (D) three types.

## Conclusion and Future Work

This project investigates SVM of ECG signal from the MIT database assuming PCA and wavelet for feature extraction at a pre-processing stage to improve classifier performance by presenting a more parsimonious input vector to the input of the classifier. Accuracies between 100% and 87% were obtained for binary and multi-class classification respectively.

The proposed approach is to be extended using adaptive wavelet where a different wavelet function is derived at each decomposing level to increase parsimony of the input vector.

Multidimensional SVM using Clifford Algebras will be used to account for multi-lead signal as well as correlated inputs from alternative sensing modalities making the proposed classification methodology relevant to future advances in sensing. The improvement of multi-class classification accuracy will be investigated as a future work using Clifford algebra [3] and Complex Support Vector Machines[4].

### Acknowledgements

I would like to thank my supervisor dr. Sillas Hadjloucas for his advice and continued support throughout this project. Moreover, I would like to acknowledge the scholarship support from the Ministry of Higher Education, Kingdom of Saudi Arabia.

### References

1. Afonso, V., Tompkins, W., Nguyen, T. and Luo, S. (1999) 'ECG beat detection using filter banks', *IEEE Transactions on Biomedical Engineering*, 46 (2), pp. 192-202.
2. Hadjloucas, S, N, Jannah, N, F Hwang, F and Galvão, R K H (2014) 'On the application of optimal wavelet filter banks for ECG signal classification', 2nd International Conference on Mathematical Modeling in Physical Sciences 2013, 240, PP.1-5
3. Bayro-Corrochano, E. J. and Arana-Daniel, N. (2010) 'Clifford support vector machines for classification, regression, and recurrence', *IEEE Transactions on Neural Networks*, 21 (11), pp. 1731-1746.
4. Boubouls, P., Theodoridis, S. and Mavroforakis, C. (2013) 'Complex Support Vector Machines for Regression and Quaternary Classification', *arXiv preprint arXiv:1303.2184*.

Detection of ECG Arrhythmia Conditions using CSVM Classifiers

N. Jannah and S. Hadjiloucas

School of Systems Engineering the University of Reading  
 n.m.h.jannah@pgr.reading.ac.uk, s.hadjiloucas@reading.ac.uk

Electrocardiogram (ECG) is widely used for the diagnosis of cardiac arrhythmia conditions. An automatic classification of four beat types Normal (N), premature ventricular contraction (PVC), Supraventricular premature or ectopic beat (SVPB) is implemented using a Multi-class Support Vector Machine (MSVM) and Complex Support Vector Machine (CSVM) algorithms [1]. The ECG signals used in these studies were obtained from the European ST-T Database. A number of beats from different leads and patients were selected for training and evaluating classifier performance. Successful ECG arrhythmia classification usually requires optimizing the following procedures: Pre-processing and beat detection, feature extraction and selection, and classifier optimization. Pre-processing and R peak detection is performed with the WFDB Software Package. This reads the annotation and finds the R (peak) location. R (peak) location used as a reference to detect peaks in other wave such P and T and extract ECG beat. ECG beats are extracted after windowing the signal using 106 samples before the R and 106 samples after the R-peak. Discrete Cosine and Sine transforms or the Discrete Fourier Transform (DFT) were used for feature extraction and dimensionality reduction of the input vector at the input of the classifier. Studies after selecting either 100 or 50 Fourier coefficients for reconstructing individual ECG beats in the feature selection phase were performed. MATLAB software routines were used to train and validate both the CSVM and the Multi-class Support Vector Machine (MSVM) classifier. A Complex kernel function, (Gaussian RBK) with 5-fold cross validation was used for adjusting the kernel values. Sequential minimal optimization (SMO) [2] was used to train the CSVM and compute the corresponding complex hyper-plane parameters. The aim of the study was to improve multi-class SVM by extending traditional SVM algorithms to complex spaces so as to simultaneously classify four types of heartbeats. Results illustrate that the proposed beat classifier is very reliable, and that it may be adopted for automatic detection of arrhythmia conditions and classification. Accuracies between 86% and 94% are obtained for MSVM and CSVM classification respectively. Using CSVM, a 4 classes problem can be classified rapidly by decomposing it into two distinct SVM tasks. Moreover, the present research confirmed that the use of selected number of Fourier coefficients to approximate the ECG beat signal and compress the input features to the classifier can lead to high classification accuracies and improve the generalization ability of the CSVM classifier. Future work on wavelet pre-processing to further compress the input space of the classifier by generating wavelets on the basis of higher order moment criteria [3] as well as alternative approaches for extending the CSVM input and output spaces to arbitrary dimension using Clifford algebra SVM [4] will be discussed at the conference.

References

- [1] P. Bouboulis, S. Theodoridis, C. Mavroforakis, and L. Evaggelatos-Dalla, "Complex Support Vector Machines for Regression and Quaternary Classification," *IEEE Trans. Neural Networks Learn IEEE Transactions on Neural Networks and Learning Systems. Syst.*, vol.26, no.6, pp. 1–15, 2014.
- [2] J. C. Platt, "Sequential minimal optimization: A fast algorithm for training support vector machines," *Adv. Kernel Methods Support Vector Learn.*, vol. 208, pp. 1–21, 1998.
- [3] S. Hadjiloucas, N. Jannah, F. Hwang, and R. K. H. Galvão, "On the application of optimal wavelet filter banks for ECG signal classification," *J. Phys. Conf. Ser.*, vol. 490, p. 012142, 2014.
- [4] E. J. Bayro-Corrochano and N. Arana-Daniel, "Clifford support vector machines for classification, regression, and recurrence," *IEEE Trans. Neural Networks*, vol. 21, no. 11, pp. 1731–1746, 2010.

Feature Extraction	Classification Algorithms
<p>Discrete Cosine Transform (DCT)</p> $X(k) = \sum_{n=0}^{N-1} x(n) \cos\left(\frac{(2k+1)\pi n}{2N}\right), k = 0, 1, \dots, N-1$ <p>Inverse DCT</p> $x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \cos\left(\frac{(2k+1)\pi n}{2N}\right), n = 0, 1, \dots, N-1$	<p>MSVM</p> $f(x) = \sum_{i=1}^M \alpha_i \langle x, \phi(x_i) \rangle + b$ <p>Support vector (if not problem)</p> $1 + \alpha \sum_{i=1}^M \frac{1}{\alpha_i} \sum_{j=1}^M \alpha_j \langle \phi(x_i), \phi(x_j) \rangle$
<p>Discrete Cosine and Sine Transforms (DCST and DST)</p> $X(k) = \sum_{n=0}^{N-1} x(n) \cos\left(\frac{(2k+1)\pi n}{2N}\right), k = 0, 1, \dots, N-1$ $Y(k) = \sum_{n=0}^{N-1} y(n) \sin\left(\frac{(2k+1)\pi n}{2N}\right), k = 0, 1, \dots, N-1$	<p>CSVM</p> $g(x) = \text{Re}\left\{ \sum_{i=1}^M \alpha_i \langle x, \phi(x_i) \rangle + b \right\}$ <p>Two hypothesis algorithms (real and complex support vector)</p> $1 + \alpha \sum_{i=1}^M \frac{1}{\alpha_i} \sum_{j=1}^M \alpha_j \langle \phi(x_i), \phi(x_j) \rangle$
<p>Wavelet filter bank for feature extraction that decomposes the input signal into different frequency bands.</p> <p>Fourier transform</p> $X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$ <p>Inverse Fourier transform</p> $x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{j\omega t} d\omega$ <p>Discrete Fourier Transform (DFT)</p> $X(k) = \sum_{n=0}^{N-1} x(n) e^{-j2\pi kn/N}$ <p>Inverse DFT</p> $x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{j2\pi kn/N}$	<p>Wavelet transform for feature extraction of peak locations</p> $W(a, b) = \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) \frac{dt}{ a }$ <p>Wavelet transform for feature extraction of peak locations</p> $W(a, b) = \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) \frac{dt}{ a }$ <p>Wavelet transform for feature extraction of peak locations</p> $W(a, b) = \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) \frac{dt}{ a }$ <p>Wavelet transform for feature extraction of peak locations</p> $W(a, b) = \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t-b}{a}\right) \frac{dt}{ a }$

ECG beat number	Classification methodology	ECG coefficient number used	Classification accuracy
400	MSVM	100	88
600	CSVM	100	94



**Abstract**

- An automatic classification of four beat types Normal (N), premature ventricular contraction (PVC), Supraventricular premature or ectopic beat (SVPB) and Fusion of ventricular and normal beat (FUSION) is implemented using a Multi-class Support Vector Machine (MSVM) and Complex Support Vector Machine (CSVM) algorithm.
- Pre-processing and R peak detection is performed with the WFDB Software Package which reads the annotation and finds the R (peak) location. R (peak) location is used as a reference to detect peaks in other wave such P and T and extract ECG beat.
- Discrete Cosine and Sine transforms or Discrete Fourier Transform (DFT) were used for feature extraction and dimensionality reduction of the input vector at the input of the classifier.
- The aim of the study was to improve multi-class SVM by extending traditional SVM algorithms to complex spaces so as to simultaneously classify four types of heartbeats. The approach can also account for phase changes in the signal.
- Accuracy between 86% and 94% are obtained for MSVM and CSVM classification respectively.
- Results illustrate that the proposed beat classifier is very reliable, and that it may be a useful tool for detection arrhythmia conditions and perform classification automatically.

**Introduction**

- ECGs provide a graphic representation of the electrical activity of the heart muscle.
- This study evaluates ECG classification using pre-processing routines, feature extraction and selection, and ECG beat classification using MSVM and CSVM.
- DCT and DST coefficients were used to extract features presented at the input vector of the MSVM classifier whereas DFT coefficients are used for creating an input vector to the CSVM classifier.
- The MSVM and CSVM classifier is used to distinguish the four ECG arrhythmias types using the corresponding complex hyper-plane parameters.
- Sequential minimal optimization (SMO) approach is used to train the CSVM and compute the corresponding complex hyper-plane parameters.

**Support Vector Machine (SVM) algorithm**

- Lagrangian function evaluation based on the dual problem:
 
$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b$$

$$L = \max_{\alpha_i} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

**Complex Support Vector Machine (CSVM) algorithm**

$$g(z) = \text{sign} \left( \frac{2}{i} \sum_{n=1}^N (a_n d_n^* + b_n d_n) r_C^*(z_n, z) + c' + i c'' \right)$$

Two hyperplanes are now defined instead of one defined using standard SVM, real and imaginary respectively:

$$\maximize_a \sum_{n=1}^N a_n - \sum_{n=1}^N a_n a_n d_n^* d_n^* r_C^*(z_n, z_n)$$

$$\maximize_b \sum_{n=1}^N b_n - \sum_{n=1}^N b_n b_n d_n^* d_n^* r_C^*(z_n, z_n)$$

Figure (1) complex hyper-planes in CSVM

**Method**

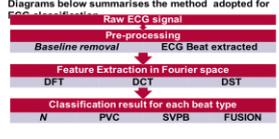


Figure (2) Structure of methodology used for ECG classification.

**Feature Extraction Algorithm**

- Reconstruction after deleting coefficient below threshold.
  - Discrete Fourier Transform (DFT) and Inverse DFT
 
$$\hat{X}(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk}, k = 0, \dots, N-1$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{X}(k) W_N^{-nk}, k = 0, \dots, N-1$$
  - Discrete Cosine and Sine transforms
 
$$C(k) = \sum_{n=0}^{N-1} x(n) \cos \left( \frac{\pi k (2n+1)}{2N} \right), k = 0, 1, \dots, N-1$$

$$S(k) = \sum_{n=0}^{N-1} x(n) \sin \left( \frac{\pi (n+1)(k+1)}{2N} \right), k = 0, 1, \dots, N-1$$
- Using DCT coefficients for creating an input vector to MSVM provides better classification than DST.

**Experimental Design and Results**

- Signal processing was carried out using MATLAB and the Wave Form Database (WFDB) software packages.
- The ECG beats were obtained from the European ST-T Database with 212 samples around the R-peak.
- In feature selection 100 Fourier coefficients were selected for reconstructing individual ECG beats, these form the input vector to the classifier shown in figure (3).
- For the classifier implementation, after feature selection is performed, the datasets are divided into two groups with 311 beats for training and testing purposes.
- The optimal hyper-plane parameters are identified on the basis of the training data set, class label. The output of the SMO algorithm was used for training.
- Finally, on the basis of the training results, test data are imported to the CSVM and MSVM classifiers to perform unknown beat classification.
- The proposed algorithm confirmed classification accuracies of 94%, whereas multi-class SVM 86% accuracies.
- Table (3) illustrates the performance of the classification process using three common measures- sensitivity (ST), specificity(SP), and predictivity (PP).

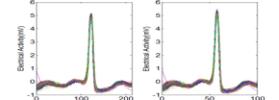


Figure (3) Comparison between original beat feature and reconstructed basis from 100 Fourier coefficients

Annotation	NORMAL	PVC	SVPB	FUSION
NORMAL	150	0	0	0
PVC	0	53	3	4
SVPB	0	0	60	0
FUSION	0	3	3	30

Table (1) confusion matrix using CSVM

Annotation	NORMAL	PVC	SVPB	FUSION
NORMAL	149	1	0	0
PVC	0	59	0	1
SVPB	0	6	54	0
FUSION	0	33	1	7

Table (2) confusion matrix using MSVM

Annotation	Classification performance result					
	CSVM		MSVM			
	PP (%)	ST (%)	SP (%)	PP (%)	ST (%)	SP (%)
NORMAL	98	100	98	100	99	100
PVC	95	88	99	60	98	84
SVPB	88	100	97	98	90	100
FUSION	88	73	99	88	17	100

Table (3) Collective result performance analysis and classification result using MSVM and CSVM

**Future Work**

- Multidimensional SVM using Clifford Algebras will be used to account for multi-lead signal analysis.
- The proposed approach is to be extended using adaptive wavelet where a different wavelet function is derived at each decomposing level to increase parimony of the input vector.
- For feature selection wavelet pre-processing will be used to further compress the input to the classifier by generating wavelets on the basis of higher order moment criteria.
- It proves convenient to write the transform in matrix form as:
 
$$t_{1 \times J} = x_{1 \times J} V_{J \times J}$$
 where  $X = [x_1, \dots, x_J]$  is the row vector of original variables,  $t$  is the row vector of new (transformed) variables and  $V$  is the matrix of weights.
- Let  $\{h_1, \dots, h_{2^{2L}}\}$  and  $\{g_1, \dots, g_{2^{2L}}\}$  be the impulse responses of the low-pass and high-pass filters respectively.
- Choosing  $v$  to be unitary:
 
$$\begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$
- The approximation  $c$  and detail  $d$  coefficients are stacked in vector  $t = [c \ 0] \ 0 \ 0$ , with coefficients  $t_j$  in larger scale associated with broad features in the data vector, and coefficients in smaller scales associated sharp peaks.
 
$$t_j = \sum_{n=0}^{2^j-1} x_n v_j(n), j = 0, 1, \dots, J-1$$

$$2^{N-1-2j} \sum_{n=0}^{2^j-1} h_n h_{n+2^j} = \begin{cases} 1, & j=0 \\ 0, & 0 < j < N \\ -1, & j=N-1 \end{cases}, n=0, 1, \dots, 2N-1$$
- Details of the optimization of the wavelet filter band parametrization are discussed in [1].

**Summary**

- Using CSVM, a 4 classes problem can be classified rapidly by decomposing it into two distinct SVM tasks.
- The present research confirmed that the use of selected number of Fourier coefficients to approximate the ECG beat signal and further compress the input features to the classifier can lead to high classification accuracies and improve the generalization ability of the CSVM classifier.
- Using the CSVM algorithm indicated a significant improvement in multi-class classification over MSVM.

**Acknowledgements**

N. Jannah acknowledges scholarship support from the Ministry of Higher Education, Kingdom of Saudi Arabia.

**References**

[1]. S. Hadjioucas, N. Jannah, F. Hwang, and R. K. H. Galvão, "On the application of optimal wavelet filter banks for ECG signal classification," *J. Phys. Conf. Ser.*, 490, 012142, 2014.



# ECG Arrhythmia Classifiers : Comparison between Multi-class SVM and CSVM Algorithms

University of Reading  
School of Systems Engineering

Najlaja Jannah and Sillas Hadjiloucas

## Introduction

According to public health agencies cardiovascular diseases are currently one of the biggest causes of death in developed countries and cardiac failure incidents are increasing every year. Electrocardiogram (ECG) analysis is one of the most commonly used tools to test and diagnose heart problems because ECGs provide a graphic representation of the electrical activity of the heart muscle. Moreover, the analysis of the ECG signals provides comprehensive information that can be used to analyse different heart conditions. This study evaluates ECG classification using pre-processing routines, feature extraction and selection, and ECG beat classification using MSVM and CSVM. An automatic classification of four beat types normal beat (N), premature ventricular contraction (PVC), Atrial premature contraction (APC) and Fusion of ventricular and normal beat (FUSION) is implemented using a Multi-class Support Vector Machine (MSVM) and a Complex Support Vector Machine (CSVM) algorithm. The aim of the study was to improve multi-class SVM by extending traditional SVM algorithms to complex spaces so as to simultaneously classify four types of heartbeats. The MSVM and CSVM classifier are used to distinguish the four ECG arrhythmias types using the DCT, DST and DFT coefficients as feature vectors at the input vector of the classifier. A sequential minimal optimization (SMO) approach is used to train the CSVM and compute the corresponding complex hyper-plane parameters.

## Non-linear Support Vector Machine (SVM) algorithm

Using a linear SVM is not possible for obtaining accurate classification results due to the data not being linearly separable. In this case, a suitable mapping function needs to be used to transform the input data to a higher dimensional feature space where it can be linearly separable as required. A general non-linear SVM can be expressed as:

$$f(x) = \sum_{i=1}^N \alpha_i y_i K(x_i, x_j) + b$$

The Lagrangian function evaluation based on the dual problem:

$$L = \max \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

## Complex Support Vector Machine (CSVM) algorithm

The adopted formulation for the current study follows the general ideas discussed by Bouboulis *et al.*, (2015) which generalised the SVM formulation to complex spaces. The approach defines a complex space separated into four parts by using a pair of complex hyperplanes as shown in figure (1). The equation below represents a decision function which is used to find the output of the classifier:

$$g(z) = \text{sign} \left( 2 \sum_{n=1}^N (a_n d_n^r + i b_n d_n^i) \kappa_C^r(z_n, z) + e^r + i e^i \right)$$

The dual problem in CSVM can be divided into two separate maximisation tasks, as there are two hyperplanes - real and imaginary:

$$\text{maximize}_a \sum_{n=1}^N a_n - \sum_{n,m=1}^N a_n a_m d_n^r d_m^r \kappa_C^r(z_m, z_n)$$

$$\text{maximize}_b \sum_{n=1}^N b_n - \sum_{n,m=1}^N b_n b_m d_n^i d_m^i \kappa_C^i(z_m, z_n)$$

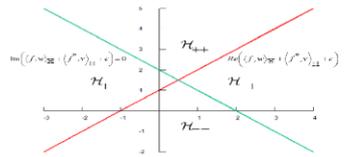


Figure (1) complex hyper-plane in CSVM

## Feature Extraction Algorithm

Reconstruction after deleting coefficient below threshold.

- Discrete Fourier Transform (DFT) and Inverse DFT

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{-nk}, k = 0, \dots, N-1$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-nk}, k = 0, \dots, N-1$$

- Discrete Cosine and Sine transforms

$$C(k) = \sum_{n=0}^{N-1} x(n) \cos\left(\frac{\pi k(2n+1)}{2N}\right), k = 0, 1, \dots, N-1$$

$$S(k) = \sum_{n=0}^{N-1} x(n) \sin\left(\frac{\pi(n+1)(k+1)}{N+1}\right), k = 0, 1, \dots, N-1$$

## Method

Diagrams below summarises the method adopted for ECG classification.

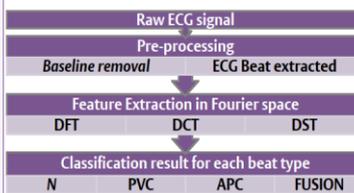


Figure (2) Structure of methodology used for ECG classification.

## Experimental Design and Results

The ECG signals used in these studies were obtained from the MIT-BIH cardiac arrhythmia database using limb lead (II) from two patient records (record 208 and record 209). These were selected for training and evaluation of classifier performance. Wave (P, QRS complex, T) detection is performed using Multiresolution Wavelet Analysis. Signal processing was carried out using MATLAB. The ECG beats were obtained with 300 samples around the R-peak. In feature selection 50 Fourier coefficients were selected for reconstructing individual ECG beats, these form the input vector to the classifier shown in figure (3). For the classifier implementation, after feature selection is performed, the datasets are divided into two groups with 500 beats for training and testing purposes. The optimal hyper-plane parameters are identified on the basis of the training data set, class label. The output of the SMO algorithm was used for training. Finally, on the basis of the training results, test data are imported to the CSVM and MSVM classifiers to perform unknown beat classification. The proposed algorithm achieved classification accuracies up to 97%, whereas multi-class SVM achieved up to 83% accuracies. Table (3) illustrates the performance of the classification process using three common measures- sensitivity (ST), specificity (SP), and predictively (PP).

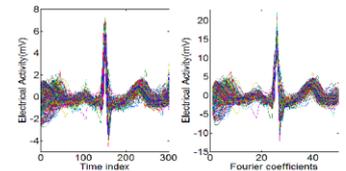


Figure (3) Comparison between original beat feature and reconstructed beats from 50 Fourier coefficients

Annotation	Output result			
	NORMAL	PVC	APC	FUSION
NORMAL	190	2	0	1
PVC	0	95	0	0
APC	1	0	100	0
FUSION	9	3	0	99

Table (1) confusion matrix using CSVM

Annotation	Output result			
	NORMAL	PVC	APC	FUSION
NORMAL	178	0	0	0
PVC	0	65	0	0
APC	20	33	100	30
FUSION	2	2	0	70

Table (2) confusion matrix using MSVM

Annotation	Classification performance result					
	CSVM			MSVM		
	PP (%)	ST (%)	SP (%)	PP (%)	ST (%)	SP (%)
NORMAL	98	95	99	100	89	100
PVC	99	95	100	100	65	100
APC	99	100	100	55	100	79
FUSION	89	99	97	95	70	99

Table (3) Collective result performance analysis and classification result using MSVM and CSVM

## Conclusion and Future Work

Using CSVM, a 4 classes problem can be classified rapidly by decomposing it into two distinct SVM tasks. The present research confirmed that the use of selected number of Fourier coefficients to approximate the ECG beat signal and further compress the input features to the classifier can lead to high classification accuracies and improve the generalization ability of the CSVM classifier. Accuracy between 83% and 97% are obtained for MSVM and CSVM classification respectively. Using the CSVM algorithm indicated a significant improvement in multi-class classification over MSVM.

In future work Multidimensional SVM using Clifford Algebras (Bayro-Corrochano and Arana-Daniel, 2010) will be used to account for multi-lead signal analysis. The proposed approach is to be extended using adaptive wavelets where a different wavelet function is derived at each decomposing level to increase parsimony of the input vector. For feature selection, wavelet pre-processing will be used to further compress the input to the classifier by generating wavelets on the basis of higher order moment criteria.

## Acknowledgements

I would like to thank my supervisor Dr. Sillas Hadjiloucas for his advice and continued support throughout this project and acknowledge the scholarship support from the Ministry education of Kingdom Saudi Arabia.

## References

Bouboulis, P., Theodoridis, S., Mavroforakis, C. and Evagelatou-Dalla, L. (2015) Complex Support Vector Machines for Regression and Quaternary Classification. IEEE Transactions on Neural Networks and Learning Systems, 26(6), 1260 - 1274.

Bayro-Corrochano, E. J. and Arana-Daniel, N. (2010) Clifford support vector machines for classification, regression, and recurrence. IEEE Transactions on Neural Networks, 21(11), 1731-1746.

# A novel technique for classifier cardiac arrhythmia condition using SIMCA

N. Jannah | S. Hadjiloucas |

## Introduction

This study investigated the use of a new ECG arrhythmia classification scheme based on a Principal Component Analysis (PCA) for feature extraction and a Soft Independent Modelling of Class Analogy (SIMCA) classifier for the automatic classification of six beat types presented simultaneously to the classifier : normal beat (NORMAL), premature ventricular contraction (PVC), atrial premature contraction (APC), right Bundle Branch Block Beat (RBBB), left Bundle Branch Block Beat (LBBB) and fusion of ventricular and normal beat (FUSION) were taken from the MIT-BIH arrhythmia database. The classification followed the time domain feature extraction. The aim of the study was to classify these six types of beats that have different dimensions simultaneously while improving the accuracy of the multi-class classification. Pre-processing of the signal and beat detection, feature extraction and compression, and classifier implementation are the three most important steps associated with successful ECG arrhythmia classification software. PCA is used to find the PCs for each beat and to create feature vectors based on the PCs to be presented at the input stage of the SIMCA classifier. Finally, the SIMCA classifier is employed to create models of each class individually and to classify the six types of ECG beats simultaneously.

## Method and Procedure

The diagrams below summarises the algorithm required in each step of performing ECG classification.

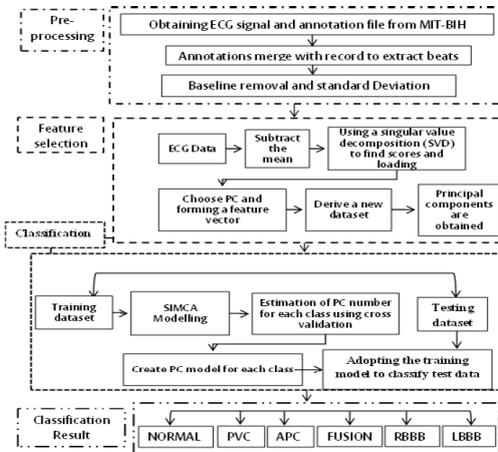


Figure 1. Block diagram of proposed method for cardiac arrhythmia classification

## SIMCA classifier

In SIMCA, the class models are developed using PCA. A cross-validation procedure is used to establish the number of PCs to be used as feature vector at the input of the classifier. A primary model is developed and one PC at a time is retained. This model is utilised in order to fit the samples into the test set and to calculate the residuals of the fit. The variance of the residuals of the sample when fit for this class is calculated from:

$$S_i^2 = \sum_{j=1}^R \frac{(e_{ij})^2}{R-K} \quad \sigma_0^2 = \sum_{i=1}^N \sum_{j=1}^R \frac{(e_{ij})^2}{(R-K)(N-K-1)}$$

where  $S_i^2$  is the variance in the residuals for the sample fit to the class,  $\sigma_0^2$  is the variance in the residuals for the class,  $e_{ij}$  are the residuals,  $R$  is the dimensionality of the class,  $K$  is the best number of PCs retained in the model, and  $N$  is the number of samples in the class [2].

## Experimental design and results

The ECG signals used in this study were based on limb lead (II) and chest lead (V1) datasets containing 2400 beats representing six ECG beat types. Each dataset includes 400 NORMAL, 400 PVC, 400 APC, 400 RBBB, 400 LBBB, and 400 FUSION segments. These were used for training and evaluation of classifier performance. Each beat relates to a segment that included 127 data point before and 128 data point after the R event, resulting in 256 points. With regard to feature selection, between 10 and 30 PC coefficients were selected to reconstruct individual ECG beats, which formed the input vector for the classifier. For the classifier implementation, after the feature selection was performed, the datasets were divided into two groups containing 1200 beats for training and testing purposes. The PCA was performed first on each group of classes and the SIMCA was then used for PCA classes. The performance of the SIMCA was based on creating individual models for each class and finding similarities between test objects and class models. These models were combined with the test dataset to assess the classification performance attained. Table 1 illustrates the performance of the classification process using four common measures- sensitivity (ST), specificity (SP), positive predictively (PP) and accuracy (ACC).

Annotation	Classification performance Output result							
	Limb lead				Chest lead			
	PP (%)	ST (%)	SP (%)	ACC (%)	PP (%)	ST (%)	SP (%)	ACC (%)
NORMAL	88.24	97.50	97.40	97.42	100	78.50	100	96.42
PVC	95.98	95.50	99.20	98.58	90.95	95.50	98.10	97.67
APC	95.94	94.50	99.20	98.42	82.99	100	95.90	96.58
RBBB	95.70	89	99.20	97.50	90.73	93	98.10	97.25
LBBB	100	100	100	100	100	100	100	100
FUSION	100	98.50	100	99.75	99.47	93	99.90	98.75

Table 1. Collective performance analysis and classification result using SIMCA

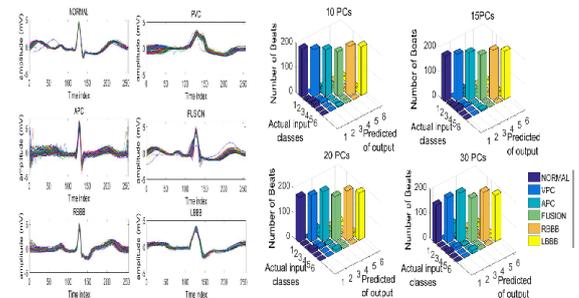


Figure 2. Extraction of six types of ECG beats and the predicted input of individual ECGs beat into the SIMCA classifier

## Conclusion and Future Work

This work demonstrated that the SIMCA classifier can be used to differentiate between the six types of ECG beats simultaneously. The 256- dimensional feature vectors around each R peak and the RR time interval were treated as inputs for the SIMCA classifier to determine six different types of ECG arrhythmias. The average classification accuracy of the proposed scheme was 98.61% and 97.78% using limb and chest lead datasets respectively. The work demonstrated the effectiveness of the proposed scheme for ECG arrhythmia classification by achieving a high accuracy ratio. In future work Multidimensional SVM using Clifford Algebras [2] will be used to classify multi-lead signals simultaneously instead of SIMCA.

### Acknowledgements

We acknowledge the scholarship support from the Ministry of Higher Education, Kingdom of Saudi Arabia.

### References

- [1] Siegel JA 2016 Forensic Chemistry: Fundamentals and Applications. USA: John Wiley & Sons, Ltd.
- [2] Bayro-Corrochano E and Arana-Daniel N 2010 Clifford support vector machines for classification, regression, and recurrence. *IEEE Transactions on Neural Networks*. 21 1731-46.

## A Comparison Between ECG Beat Classifiers Using Multiclass SVM and SIMCA with Time Domain PCA Feature Reduction

Najlaa Jannah

*School of Biological Sciences, Department of  
Biomedical Engineering  
University of Reading  
Reading, United Kingdom  
e-mail: n.m.h.jannah@pgr.reading.ac.uk*

Sillas Hadjiloucas

*School of Biological Sciences, Department of  
Biomedical Engineering  
University of Reading  
Reading, United Kingdom  
e-mail: s.hadjiloucas@reading.ac.uk*

**Abstract**—Detection and treatment of arrhythmias has become one of the main goals in cardiac care diagnosis provided by general practitioners. Electrocardiogram (ECG) analysis is one of the most commonly used tools to test and diagnose heart problems. Classification of ECG heartbeats enables the identification of specific arrhythmia or other heart conditions. This paper presents and contrasts the results from two effective ECG arrhythmia classification schemes. The first scheme consists of a principal component analysis (PCA) step for feature reduction at the input vector to the classifier, combined with soft independent modelling of class analogy (SIMCA). The second method uses a multi-class support vector machine (MSVM) classifier to differentiate between four different types of arrhythmia from ECG beats. The four types of beats include Normal (N), Premature Ventricular Contraction (PVC), and Atrial premature contraction (APC) and Right Bundle Branch Block Beat (RBBB). The time domain features were obtained from the St Petersburg INCART 12-lead Arrhythmia Database (incartdb). Between 10 and 30 Principal Components (PCs) were selected for reconstructing individual ECG beats and create the input vector to the classifier. The average classification accuracy of the proposed scheme is 76.83% and 98.33% using MSVM and SIMCA classifier respectively. The SIMCA classification algorithm provided better performance than the MSVM classifiers.

**Keywords**— *Electrocardiogram; Principal Component Analysis; Soft independent modeling of class analogy; Multi-class Support Vector Machine*

### I. INTRODUCTION

According to public health agencies cardiovascular diseases are currently one of the biggest causes of death in developed countries and cardiac failure incidents are increasing every year. ECG is widely used for the diagnosis of cardiac arrhythmia conditions as signals ECG provides comprehensive information that can be used to analysis different heart conditions. Moreover, ECG classification of heartbeats enables the identification of specific arrhythmia or other heart condition types. In recent decades, many algorithms for automatic ECG heartbeat classification have been developed. There are several accounts of ECG beat classification in the literature using different techniques such as neural networks [1], recurrent neural networks (RNN) [2],

probabilistic neural network (PNN) [3], deep learning neural network (DNN) [4], support vector machine (SVM) classifiers [5],[6] and complex SVM (CSVM) [7].

In the study reported in [2], four types of ECG beats were considered: N, congestive heart failure, ventricular tachyarrhythmia and atrial fibrillation beats, these were successfully classified using an RNN trained with the Levenberg–Marquardt algorithm. In another study, SVM in conjunction with multi-class directed acyclic graphs was used to classify four types of ECG signals (N, Atrial fibrillation, Ventricular tachyarrhythmia and Congestive heart failure). The Directed Acyclic Graph Support Vector Machine (DAGSVM) classifier has been considered as one of the important recent advances in this line of research and has yielded an average accuracy rate of 97.71% [8].

Looking at the above studies collectively, it transpires, that even though many works have been reported in the arrhythmia beat classification literature, there is still a need to improve further the classification accuracy by creating a model for each class individually while accounting in the classification algorithm for different beats duration. A comparison of the performance of ECG beats classification using single- and multi-model approaches merits further investigation so as to develop a holistic and effective treatment programme. The aim of this study is to develop and assess such new methodologies and provide a comparison of the performance of the classification offered by using single-model (MSVM) and multi-model (SIMCA).

The study focuses on the use of a new ECG arrhythmia classification scheme based on PCA for feature extraction and either an MSVM or a SIMCA classifier to differentiate between four type of arrhythmias conditions that were obtained from records in the St Petersburg INCART 12-lead Arrhythmia Database (incartdb) [9]. Pre-processing of the signal and beat detection, feature extraction and compression, and classifier implementation are the three most important steps associated with successful ECG arrhythmia classification software. PCA is used to find the principal components (PCs) associated with each beat and to create feature vectors based on the PCs that were presented at the input stage of the MSVM and SIMCA classifiers. Both the MSVM as well as the SIMCA classifier were used to classify the four types of ECG beats (annotated as Normal (N), Premature Ventricular Contraction (PVC), and Atrial