On Designing an ECG-based Intelligent System: Utilizing the Heart’s Electrical Activity to Recognize Humans and Detect Arrhythmia

Sara Saeed Abdeldayem
West Virginia University, ssa0006@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Computer Engineering Commons, and the Signal Processing Commons

Recommended Citation
https://researchrepository.wvu.edu/etd/3702

This Thesis is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Graduate Theses, Dissertations, and Problem Reports by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.
ON DESIGNING AN ECG-BASED INTELLIGENT SYSTEM: UTILIZING THE HEARTS ELECTRICAL ACTIVITY TO RECOGNIZE HUMANS AND DETECT ARRHYTHMIA

Sara Saeed Abdeldayem

Thesis submitted to the
Benjamin M. Statler College of Engineering and Mineral Resources
at West Virginia University
in partial fulfillment of the requirements for the degree of
Master of Science
in
Computer Science

Thirimachos Bourlai, Ph.D., Committee Chairperson
Yuxin Liu, Ph.D.
Yangfang Ye, Ph.D.
Vinod K. Kulathumanii, Ph.D.
Lane Department of Computer Science and Electrical Engineering

Morgantown, West Virginia
2018

Keywords: Electrocardiogram (ECG), Arrhythmia, Texture, Signal Analysis, Deep Learning, Spectro-temporal, Spectral Correlation, Identification, Verification

Copyright © 2018 Sara Saeed Abdeldayem
ABSTRACT

On Designing an ECG-based Intelligent System: Utilizing the Hearts Electrical Activity to Recognize Humans and Detect Arrhythmia

Sara Saeed Abdeldayem

The electrocardiogram (ECG) signal is the bioelectrical signal that reflects the heart’s activity. It has been extensively used as a diagnostic tool since it holds information about the cardiac health condition. However, recent researches have shown that it exhibits an inter-subject variability property. Therefore, it can be used as a biometric-based modality for either identification or verification purposes.

Nevertheless, some of the challenges are faced while employing such a signal. For instance, ECG signal is prone to noise, accordingly, noise filters should be designed to remove the noise while keeping the signal properties. Moreover, factors such as medications, health condition, and emotional state can affect the intra-subject variability of the signal. Therefore, experiments that are performed should consider these factors to produce a robust system.

In this thesis, we utilize the ECG signals to propose three approaches: one clinical approach, and two biometric approaches. In the first approach, we automatically detect cardiac arrhythmia, a disease when the heart fails to contract at the normal rhythm, using the signal textures. These texture operators include both temporal and spectro-temporal features. The temporal texture includes the one-dimensional local binary pattern (1D LBP), while the spectro-temporal textures include the investigation of the texture of short-time Fourier transform (STFT) and generalized Morse wavelet transform (CWT). Different classifiers as well as different ECG lead configurations were studied. Classification accuracies of 92.97%, 99.24%, and 99.81% on MITDB were achieved using the 1D LBP, CWT, and STFT textures respectively.

On the other hand, the ECG signal showed some advantages over other biometric modalities such as face and iris. These advantages incorporate the embedded liveness detection, the difficulty to counterfeit when compared to other traits, and the hybrid information it holds about the subject identity as well as the health, mental, and emotional conditions.

In this work, we propose two biometric approaches. The first approach utilizes the spectro-temporal properties of the signal for both subject identification and validation. A two-dimensional convolutional neural network (2D CNN) was employed to extract the high-level features to distinguish subjects. Both the STFT and CWT were investigated as features. Validating on eight databases, an overall identification rate of 97.85%, equal error rate of 0.0268, and area under curve of 0.99% were achieved for STFT features.

The second approach utilizes the cyclostationary properties of the signal since the ECG signal is non-stationary. The superiority of this approach is that it follows the blind-segmentation where no ECG wave points are detected, as well as being more robust to the noise in the ECG signal. The spectral function of the signal is extracted and used as features to feed a 2D CNN. Seven databases were used to validate the robustness and generalization of the approach, where an average identification rate of 94.8%, false rejection rate of 0.002% and false acceptance rate of 0.04% were achieved across all the databases.

Though, all the proposed algorithms showed better performance when compared to the state of art in the literature.
I dedicate this thesis to everyone who got forgotten in the abyss of life, either under the ground or still above...
First and foremost, I would like to thank God for being here with me all the time, in ups and downs, in happiness and sadness, and in joy and misery. I would like to thank Prof. Thirimachos Bourlai for his constant support, providing a good environment to work, and all the help that he provided during this journey. I would like also to thank the committee members for their feedback and support: Dr. Yuxin Liu, Dr. Yangfang Ye, and Dr. Vinod K. Kulathumanii.

Special thanks for my family for being my all-time supporters: my mom, my sister, and my brothers, and life time friend Eman Hosny. For always trying to make me happy despite the problems that they have on their own.

I would also like to thank Fulbright Foreign Student Program for giving me such an opportunity to widen my horizon, meet new people, and merge with a new culture.

Kudos to my lab squad: Jake Rose, Erin Moore, and Suha Reddy, as well as the stuff of LCSEE department for their assistance during those two years.

Finally, thanks to the Global Intervarsity and the WVU Collegiate Recovery center and everyone who makes me feel home and for all the dinners, tailgates, yoga classes, and the events that I participated in.
Contents

Abstract ii

Acknowledgments iv

List of Figures viii

List of Tables x

1 Introduction 1
  1.1 Clinical ECG . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
  1.2 Biometric ECG . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
  1.3 Challenges . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
  1.4 Contributions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
    1.4.1 Automatically detecting Arrhythmia in ECG Signals . . . . . . . . . . . . . . . 5
    1.4.2 ECG-based Individual Identification . . . . . . . . . . . . . . . . . . . . . . . . 6
  1.5 Organization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
  1.6 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8

2 Medical Background 9
  2.1 Heart Electrical Activity . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 9
  2.2 Recording the ECG signal . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
  2.3 ECG Abnormalities . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
  2.4 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

3 Literature Review 18
  3.1 ECG in Diagnosis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
    3.1.1 Time Domain Features . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
3.1.2 Frequency Domain Features .................................................. 20
3.1.3 Time-frequency Domain Features .......................................... 20
3.2 ECG in Biometrics ................................................................. 21
  3.2.1 Fiducial Features ............................................................... 21
  3.2.2 Non-fiducial Features ......................................................... 23
  3.2.3 Hybrid Features ................................................................. 26
3.3 Summary ................................................................. 27

4 Methodology ................................................................. 28
  4.1 ECG Databases ................................................................. 28
  4.2 ECG Signal Processing ........................................................ 31
    4.2.1 Noise Removal ............................................................... 31
    4.2.2 QRS Detection .............................................................. 32
  4.3 ECG Features ................................................................. 33
    4.3.1 1D Local Binary Pattern (LBP) ........................................... 34
    4.3.2 Time-Frequency Features ................................................. 35
    4.3.3 Cyclostationary Signals .................................................... 38
  4.4 Classifiers ................................................................. 42
    4.4.1 Support Vector Machines (SVM) .......................................... 43
    4.4.2 Deep Convolutional Neural Network (CNN) .......................... 45
  4.5 Performance Metrics ........................................................ 47
  4.6 Summary ................................................................. 48

5 Automatic Detection of Arrhythmia ........................................ 49
  5.1 Introduction ................................................................. 49
  5.2 Experimental Setup ........................................................ 50
    5.2.1 Signal Preprocessing ...................................................... 50
    5.2.2 Beats Extraction ............................................................ 50
    5.2.3 Feature Extraction .......................................................... 50
    5.2.4 Classification and model validation ..................................... 52
    5.2.5 Lead Configurations ....................................................... 52
  5.3 Results ................................................................. 53
  5.4 Summary ................................................................. 57
6  ECG-based Human Authentication using High-level Spectro-temporal Features  58

6.1  Introduction ................................................................. 58

6.2  Experimental Setup ....................................................... 59

6.2.1  Signal Preprocessing ................................................... 59

6.2.2  Feature Extraction ....................................................... 59

6.2.3  Individual Identification and Verification ........................... 60

6.2.4  Model Validation .......................................................... 61

6.3  Results ........................................................................... 62

6.4  Summary ........................................................................ 65

7  ECG-based Human Identification using Spectral Correlation  66

7.1  Introduction .................................................................... 66

7.2  Experimental Setup ....................................................... 67

7.2.1  Signal Preprocessing ................................................... 67

7.2.2  Feature Extraction ....................................................... 67

7.2.3  Individual Identification ................................................ 67

7.2.4  Model Validation .......................................................... 68

7.3  Experimental Results ..................................................... 68

7.4  Summary ........................................................................ 72

8  Discussion and Conclusions .................................................. 73

Bibliography ........................................................................ 74
## List of Figures

1.1 Intra-subject variability of ECG heartbeats, from three randomly selected subjects from of the MIT-BIH database (a) Normal beats. (b) Right block bundle beats. (c) Left block bundle beats. ........................................... 5

2.1 Human heart conduction system [1] .................................................. 10
2.2 Cardiac heartbeat trace and wave points. ........................................ 11
2.3 Cardiac events of the ECG beat [1]. ............................................... 12
2.4 The limb leads and the augmented limb leads in ECG recording system. .... 13
2.5 The precordial leads in ECG recording system. ................................. 14
2.6 The shape of one heartbeat in different ECG lead [2]. ......................... 15
2.7 Interpretation of left and right bundle branch block diseases from ECG signals of different leads. [3]. ............................................................. 17

4.1 Examples of the noisy ECG and corresponding filtered signal. ............... 32
4.2 The general workflow of the QRS detector. ...................................... 33
4.3 The effect of generalized Morse wavelet parameters on the wavelet shape [4]. 35
4.4 The main idea of short-time Fourier transform. ................................. 36
4.5 Hamming window function and its corresponding Fourier transform [5]. .... 37
4.6 The effect of generalized Morse wavelet parameters on the wavelet shape [6]. ......................... 38
4.7 Time instants $t_1$ and $t_2$ separated by a lag $\tau$. ............................ 40
4.8 The cyclostationary function of a synthetic signal [7]. ....................... 42
4.9 Support Vector Machine (SVM) data separation. ............................ 43
4.10 SVM margin for training set $\mathcal{S}$ ............................................ 44
4.11 Neural Network layers architecture [8]. ........................................... 46

5.1 Workflow of the arrhythmia detection algorithm. ............................ 50
5.2 The STFT images of the four beats types. The second row is the difference between each heartbeat type image and the normal image. The third row is the GLCM segments images for each beat [4].

5.3 The CWT images of the four beats types. The second row is the difference between each heartbeat type image and the normal image. The third row is the GLCM segments images for each beat [4].

5.4 R peak detection in the ECG signal of one of the records in MITDB.

5.5 Effect of P size on the testing data for both classifiers using the two ECG leads, where the SVM performance is in solid, and KNN is in dotted lines [4].

5.6 Accuracy results after applying cross-validation for all lead configurations using SVM classifier [4].

5.7 Performance of each ECG beat type using different configurations and SVM classifier [4].

6.1 Workflow of the ECG biometric approach using high-level spectro-temporal features.

6.2 The intra- and inter-subject variability of both STFT (left) and CWT (right) images of two subjects from the MITDB [9].

6.3 The architecture of the proposed CNN using the spectral image as input and the extracted feature to be compared with other subjects for identification.

6.4 The CMC curve of the proposed approach of both STFT and CWT using ECG different databases [9].

6.5 The equal error rate (EER), and area under curve (AUC) of both proposed approaches on the eight ECG databases.

6.6 The receiver operating curves of both STFT and CWT features using the proposed approach.

7.1 Workflow of the ECG biometric approach using the cyclostationary properties.

7.2 The intra- and inter-subject variability in the spectral correlation images of averaged 10 ECG segments. Each row represents two different groups of 10-averaged ECG segment per one subject and their difference to show the intra-subject variability. Whereas columns represent the difference between the images between different subjects in MITDB database.

7.3 The boxplot of 10 validation experiments on each dataset using the proposed approach.

7.4 The CMC curve of the proposed approach on the validation databases.
List of Tables

2.1 ECG electrodes acronyms and positions. ........................................... 13

4.1 ECG databases in the literature work. ............................................. 29

5.1 Distribution of training and testing data classes. ............................... 53

5.2 LBP texture feature performance on the MITDB dataset. ................. 54

5.3 STFT texture feature performance on the MITDB dataset. ............... 55

5.4 CWT texture feature performance on the MITDB dataset. ............... 55

5.5 Comparison of the proposed approach with intra-patient approaches in the literature using the MITDB database. ................................. 57

6.1 The proposed CNN layers information. .......................................... 61

6.2 The performance of the ECG databases using the proposed system. .... 64

6.3 Comparison of the average performance of proposed approach with the literature work. 64

7.1 The proposed approach’s average identification rates on different ECG databases. . 70

7.2 Comparison of the proposed approach with the literature work on the databases level. 71

7.3 Comparison of the average performance of proposed approach with the literature work. 72
<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACF</td>
<td>Autocorrelation Function</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>AUC</td>
<td>Area Under Curve</td>
</tr>
<tr>
<td>CAF</td>
<td>Cyclic Autocorrelation Function</td>
</tr>
<tr>
<td>CF</td>
<td>Cyclic Frequency</td>
</tr>
<tr>
<td>CMC</td>
<td>Cumulative Match Characteristic</td>
</tr>
<tr>
<td>CNN</td>
<td>Convolutional Neural Network</td>
</tr>
<tr>
<td>CWT</td>
<td>Continuous Wavelet Transform</td>
</tr>
<tr>
<td>DCT</td>
<td>Discrete Cosine Transform</td>
</tr>
<tr>
<td>DWT</td>
<td>Discrete Wavelet Transform</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiogram</td>
</tr>
<tr>
<td>EER</td>
<td>Equal Error Rate</td>
</tr>
<tr>
<td>FAR</td>
<td>False Acceptance Rate</td>
</tr>
<tr>
<td>FC</td>
<td>Fully Connected Layer</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FN</td>
<td>False Negative</td>
</tr>
<tr>
<td>FP</td>
<td>False Positive</td>
</tr>
<tr>
<td>FRR</td>
<td>False Rejection Rate</td>
</tr>
<tr>
<td>GLCM</td>
<td>Grayscale Co-occurrence Matrix</td>
</tr>
<tr>
<td>HR</td>
<td>Heart Rate</td>
</tr>
<tr>
<td>ICA</td>
<td>Independent Component Analysis</td>
</tr>
<tr>
<td>IDR</td>
<td>Identification Rate</td>
</tr>
<tr>
<td>KNN</td>
<td>K-nearest Neighbor</td>
</tr>
<tr>
<td>LBP</td>
<td>Local Binary Patterns</td>
</tr>
<tr>
<td>RNN</td>
<td>Recurrent Neural Network</td>
</tr>
<tr>
<td>ROC</td>
<td>Receiver Operating Characteristics</td>
</tr>
<tr>
<td>SCF</td>
<td>Spectral Correlation Function</td>
</tr>
<tr>
<td>STFT</td>
<td>Short-time Fourier Transform</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>TN</td>
<td>True Negative</td>
</tr>
<tr>
<td>TP</td>
<td>True Positive</td>
</tr>
</tbody>
</table>
### Nomenclature

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LBP_p(c)$</td>
<td></td>
<td>Local binary pattern at point $c$ with neighbors $p$</td>
</tr>
<tr>
<td>$\omega$</td>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td>$\tau$</td>
<td></td>
<td>Delay</td>
</tr>
<tr>
<td>$x(t)$</td>
<td></td>
<td>Time signal</td>
</tr>
<tr>
<td>$\mathcal{F}{\cdot}$</td>
<td></td>
<td>Fourier transform</td>
</tr>
<tr>
<td>$STFT{\cdot}$</td>
<td></td>
<td>Short-time Fourier transform</td>
</tr>
<tr>
<td>$\omega(n)$</td>
<td></td>
<td>Windowing Function</td>
</tr>
<tr>
<td>$W_{\psi}$</td>
<td></td>
<td>Wavelet transform function</td>
</tr>
<tr>
<td>$\Psi_{\beta,\gamma}$</td>
<td></td>
<td>Generalized Morse wavelet function</td>
</tr>
<tr>
<td>$C_{xx}$</td>
<td></td>
<td>Autocovariance</td>
</tr>
<tr>
<td>$R_x$</td>
<td></td>
<td>Autocorrelation function</td>
</tr>
<tr>
<td>$\alpha$</td>
<td></td>
<td>Cyclic Frequency</td>
</tr>
<tr>
<td>$S^\alpha_x$</td>
<td></td>
<td>Spectral correlation function</td>
</tr>
<tr>
<td>$g(\cdot)$</td>
<td></td>
<td>Regression function</td>
</tr>
<tr>
<td>$\hat{\gamma}$</td>
<td></td>
<td>Support vector machine margin</td>
</tr>
<tr>
<td>Acc.</td>
<td></td>
<td>Accuracy</td>
</tr>
<tr>
<td>Sens.</td>
<td></td>
<td>Sensitivity</td>
</tr>
<tr>
<td>Spec.</td>
<td></td>
<td>Specificity</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The electrocardiogram (ECG or EKG) is an electrical signal that reflects the electrical activity of the cardiac cycle. It can be recorded by using electrodes that are placed on patient’s body surface. These electrodes measure the changes in the voltage that are caused by the action potentials of the excitable cardiac cells. The resulting variations forms the heartbeat, that in turn manifest a series of waves. These waves’ morphologies are timing carry information about the electrical activity of the heart and accordingly the diagnosis of the heart disease if exists [10]. Not only the morphology of the heartbeat that can hold information about the health condition of the heart, but also the time pattern between beats. For instance, to indicate the heart rate.

The ECG has been broadly used as a clinical tool to diagnose cardiac diseases studies since it contains rich information about the hearts activity. For instance, it is considered as a common tool to diagnose cardiac arrhythmia. Cardiac arrhythmia happens when the heart fails to beat or contract at the correct rhythm. It can be considered as a common cause of death and according to the cardiac disease report of 2015, it affects more than one million in the united states only. Nevertheless, some researchers have studied the utilization of the ECG signals to identify individuals in biometric systems.

In this thesis we study both aspects of the ECG; the clinical aspect and the biometric one. We propose a fully automatic arrhythmia detection algorithm as well as two ECG-based biometric approaches.

In the proceeding section we will give a brief background about the two aspects.
1.1 Clinical ECG

Since the health condition of the heart activity can be reflected on its ECG, we can utilize this signal to detect abnormal heartbeats. Therefore, a physician can use the ECG signal of a patient to detect any cardiac diseases or check the patient’s heart rate (e.g. a faster or slower than normal, or an irregular heartbeat). Afterwards, if the doctor finds any abnormality in the ECG signal, further cardiac assessment tests can be done to decide whether a treatment or monitoring is necessary.

Although not all cardiac arrhythmias are fatal, it still need monitoring as the heart condition may deteriorate if not treated soon.

The traditional arrhythmia detection approaches including manual investigation of the ECG records are time consuming and prone to error beside being sometimes subjective [11, 12]. Therefore, an alternative way is to introduce a computer-based approach that does not require the presence of a medical expert. These software programs can be installed in any computer, as the ECG signal can be loaded and then, modern signal processing and machine learning approach is utilized to determine if the signal has abnormality that needs treatment or further investigation. Therefore, it can save the doctor’s time as she/he may need to assess multiple ECG signals from many patients in a timely manner, reduce the error that may result from subjective point of views, or can be used by an unexperienced physician in rural areas (i.e. professional physicians are not always available). However, doctors will always have the final decision to assess they system’s reported outcome. Therefore, they will save the time verifying the condition instead of manually investigating it.

1.2 Biometric ECG

The advances in technology and information nowadays such as personal health and financial information have demanded the protection of access to only authorized personnel in order to prevent breaches, ensure confidentiality and integrity. This protection requirements have been of interest for thousands of years such as giving a physical key to only legitimate users. However, these physical keys are easy to be cloned or stolen.

On the other hand, other traditional procedures of limiting the access to only legitimate users have been introduced such as ID cards, tokens, or passwords. However, these procedures are also vulnerable to identification theft and circumvention. For instance, software that steal passwords, or even cameras that capture the ID card’s image or the password written. Recently, biometric-based systems have been introduced to provide only authorized subjects access.
Biometrics methods that depends on the individual’s characteristics have gained much attention. This is because it is an automatic and the user does not need to remember a password or use card lock, where attacker who correctly observed an input password can impersonate the corresponding user freely [13].

Biometric identification can be defined as the process of discriminating a set of individuals based on their either physiological or behavioral characteristics, or both. The biometric traits that used in the literature can be categorized into physiological or behavioral attributes or traits. In the physiological attributes depends on the physical characteristics of the individual’s biometric such as the shape or the anatomy, which differs from one subject to another. Several biometric approaches have been introduced in the category. For instance, face recognition has been widely studied for individual’s identification and verification [14, 15, 16, 17]. Other physiological traits such as iris [18, 19], ear pattern [20, 21], hand geometry [22, 23], and fingerprint [24] have been studied for human recognition.

Contrarily, the behavioral attributes are related to the measure of uniquely identifying patterns in human activities. For instance, gait biometrics [25], signature [26, 27], keystroke dynamics [28, 29], and speech pattern [30].

However, each biometric attribute or trait should satisfy a set of properties but at different levels [31]. These properties include:

- Universality: the trait should be possessed by each individual.
- Uniqueness: the trait should be unique for each subject. Therefore, it should to be able to distinguish individuals successfully.
- Collectability: the trait should be possible to obtain from the individuals.
- Satisfaction: the trait should be collected in a satisfactory way with respect to the subjects.
- Permanence: the trait is stable over time.

For instance, voice and face attributes suffer from a relatively low uniqueness, while signature biometrics have universality.

Some researches on the ECG signals showed that physiological and geometrical-based characteristics of the individuals can create an inter-individual distinction [32, 33, 34, 35, 36]. This distinction comes from the variations of the mass orientation and conductivity of the heart, the difference in cardiac muscle, and the cardiac activation order of each individual[37]. Moreover, geometrical vari-
ations including the position and orientation of the heart, as well as other factors may also increase the inter-individual distinction [37].

ECG biometrics offer several advantages over other biometric traits in the following manner:

(1) The ECG signals-based biometric systems are difficult to counterfeit since it is extremely difficult to be mimicked. For instance, in the iris recognition system, one can trick the system by using fake iris images or contact lenses [38].

(2) The ECG signals provide an intrinsic liveness evidence, when compared to other biometric modalities such face and iris. For instance, iris and face do not have a proof of liveness.

(3) The ECG signal holds information heart condition and health status as discussed before as well as mental and emotional status.

Nevertheless, there are some factors that affect intra-individual variability of the ECG signal, which challenge its usage as a biometric trait. These factors can either be long-term or short-term. The long-term factors may involve the age and health condition, whereas the short-term factors may include the emotional and mental conditions and medications side effects. Figure 1.1 shows the intra-subject variability of the ECG signals of a selected number of individuals (using the MIT-BIH database). For each subject, 200 beats were plotted, each beat is composed of 200 sample points or equivalently about 0.56 seconds, which is considered sufficient to capture most of the heart cycle [39]. As shown in Figure 1, the ECG signal varies due to the fact that it represents either a normal or a potential arrhythmia related condition (i.e. left and right block bundle conditions) as well as variations in the same heartbeat type, and thus, all ECG signals exhibit amplitude and shape variations. Figure 1.1 shows the intra-subject variability of the MIT-BIH database of different heart conditions. Twenty hundred ECG heartbeats from different subjects were segmented. Each heartbeat is 0.56 seconds, or 200 sample point. As shown in the figure, the ECG signal varies in shape between the normal and arrhythmia heartbeats (i.e. left and right block bundle conditions) as well as variations in the same heartbeat type for the same subject.

Therefore, based on the previous discussion, ECG is can be considered as a suitable biometric trait.

1.3 Challenges

Several challenges face the utilization of the ECG signal in both clinical and biometric application. These challenges include but are not limited to the following:
Figure 1.1: Intra-subject variability of ECG heartbeats, from three randomly selected subjects from the MIT-BIH database (a) Normal beats. (b) Right block bundle beats. (c) Left block bundle beats.

- Noise Contamination: The collected ECG signal is contaminated with noise due to the interference of powerline in the surrounding environment, muscle contraction, and respiration process. Although the noise amplitude can reach 300 mV when the signal has values of 0.5 mV, noise is usually removed using the frequency information of each source.

- ECG variability: The factors that affect the variability can either be long-term or short-term factors. The age and health condition are examples of the long-term factors, whereas the short-term factors may involve the emotional and mental conditions.

- Segment Size: The ECG segment of interest should be long enough to extract the required information that makes it distinguishable, while it should be short enough such that the collection of data is feasible.

1.4 Contributions

In this thesis we propose three different approaches. One of the proposed approaches is concerned about the detection of abnormal heartbeats, while the other two approaches are concerned about the utilization of ECG in biometric systems.

1.4.1 Automatically detecting Arrhythmia in ECG Signals

In this approach, we propose a fully automated system to detect abnormal heartbeats using the patients’ ECG signals. It utilizes the texture of the ECG signal in both the temporal and spectro-temporal domains to detect and classify four types of heartbeats. Based on the literature work and to the best of our knowledge, none of the literature methods utilized the texture in ECG signal analysis in the 2D space.
In this research, we propose a methodological approach that automatically classify different heartbeat conditions, namely paced, left bundle block, right bundle block, and normal heartbeats. The original ECG signal is first preprocessed to remove the noise. Then, the R peaks are identified, and the signals are segmented into heartbeats. Next, different texture operator were investigated: one-dimensional local binary patterns (1D LBP) in the temporal domain are utilized, a 2D LBPs and Haralick texture features extracted from the grayscale co-occurrence matrix (GLCM) are utilized in the spectro-temporal domain. We utilize both the short-time Fourier transform (STFT) and Morse wavelets in the spectral domain investigation. Finally, we studied different classifiers and lead configurations to examine which approach is more efficient.

1.4.2 ECG-based Individual Identification

We consider two novel approaches in individual identification and verification using the subject’s ECG signal.

**Human Identification using spectral correlation**

In this approach, we utilize the electrocardiogram (ECG) as a biometric modality in human identification system. The main contribution of this work is as follow:

- The cyclostationary signal’s feature has been used to distinguish signals with different modulation schemes specially in digital radio communication signals, which are usually contaminated with noise. To the best of our knowledge, no literature work utilized these properties to identify individuals. In this work, we utilize the cyclostationary property of the ECG in the proposed approach.

  By using the spectral correlation function of the ECG segments, we can enrich the feature representation of the signal and gain more information about the distinctive characteristics of the individuals.

- Instead of depending on the extraction of the fiducial ECG points (i.e. fiducial-based approaches), we utilize a fixed duration blind segmentation approach of the ECG signal. Therefore, less computational complexity can be achieved, while improving also the performance.

- Unlike the literature work that removes the noise from the ECG signal as a preprocessing step, we take advantage that the noise does not exhibit cyclostationary properties to disregard the
noise removal step from the workflow. Thus, we validate the robustness of the system to noise that accompanies the ECG signal.

- We employ deep learning, specifically 2D convolutional neural network that reduces the need for feature engineering and provide a better performance. For instance, most of literature work that utilize feature selection use linear algorithms. However, in the deep learning, the features are learned during the training process.

- Most of the work in the literature include only the healthy subjects in the experimentation or validate on few databases. Here, we validate proposed approach on seven public databases including both normal and abnormal ECG signals. Thus, the same system can be utilized to different databases without changing the parameters, which measures the generalization ability of the system.

**Human authentication using high-level spectro-temporal features**

In this approach, we utilize the spectro-temporal images of the ECG signals similar to the arrhythmia detection algorithm. However, instead of extracting the texture feature from the images, we employ deep learning to learn the extracted features. The contribution of this work is as follow:

- Away from the temporal and morphological features that were proposed in the literature, we utilize the spectro-temporal domain features to study the changes in the frequency components along the time. We utilize the short-time Fourier transform and generalized Morse wavelets and compare between their approaches.

- Instead of using classifiers output, we propose a more distinctive features by extracting the features from the network’s intermediate layers.

- We compress the size of the template vector stored in the gallery database by utilizing a vector of size 100 for an ECG signal of 200 sample point.

- We thoroughly test the proposed model robustness on eight different ECG databases. These databases include both normal and abnormal heartbeats. Therefore, we do not exclude unhealthy subjects and try to mimic the real-world scenario.
1.5 Organization

In this chapter we gave a brief introduction about the ECG signal and its application in both clinical and biometric systems. We also discussed the proposed arrhythmia detection algorithm as well as the two proposed ECG-biometric approaches.

The thesis is organized as follows:

Chapter 2 discussed the medical background including how the ECG signal is measured and the lead configurations during the recording. It also introduces brief information about the abnormalities in the cardiac cycle that cause cardiac disease which we proposed an approach to detect.

Chapter 3 introduces the literature work in both the clinical aspect of the ECG as well as the biometric-based systems.

Chapter 4 discusses the experimental set up of the approaches including the databases that we used, the signal processing step, the features that were utilized in our approaches, as well as a concise discussion about the classifiers employed.

Chapter 5 introduces the proposed arrhythmia detection approach. It also illustrated the workflow including the beats extraction, features creation and the different lead and classifiers configuration used as well as the results.

Chapter 6 considers the first ECG-biometric based approach we propose, which utilizes the high-level spectro-temporal features. It also includes the methodology followed and the results.

Chapter 7 discusses the second ECG-biometric based approach that depends on the signal spectral correlation function. It also investigates the methodology and presents the experimental results.

Chapter 8 concludes the thesis and discusses the main points noted in the results.

1.6 Summary

In this chapter we gave an introduction about the ECG applications in both medical or clinical use such as arrhythmia detection, and biometric-based system for individual identification. We also analyzed the challenges that can be faced during the utilization of the ECG signal. Moreover, we discussed the proposed approaches in this thesis as well as our contribution in each approach. In the next chapter, we will explore the cardiac cycle, how to obtain the ECG signal, and the abnormalities that can be found in the ECG signal.
Chapter 2

Medical Background

The history of the electrocardiogram started in 1842 when Carlo Matteucci, a professor of physics at the University of Pisa, noticed an electrical current accompanies the heartbeats in a frog. Later, in 1887, Augustus Waller, a British physiologist of St Mary’s Medical School in London, made the first ECG recording on human using capillary electrometer and chest electrodes. Afterwards, in 1901, William Einthoven developed the ECG recording and used the built string-galvanometer-based three-lead ECG machine.

Before proceeding with the electrical activity, we shall give a brief anatomical and functional background about the heart.

2.1 Heart Electrical Activity

The heart is a muscular organ whose primary function is to provide the body with an oxygen-rich blood. The anatomy in a simple way can be explained as four champers, two atria and two ventricles, where we have one atrium and one ventricle in each side (i.e. left and right). The atrium role is to receive the blood, while the ventricle is to push the blood into further circulation. However, all the four champers work together in a synchronized rhythmic fashion.

The blood starts its journey by entering the right atrium in a deoxygenated form. Then, it moves to the right ventricle that pushes it to the lungs to get oxygenated through the Pulmonary arteries. The blood is then filled with oxygen from the lungs and streamed back to the left atrium though the pulmonary vein. The left atrium then passes the oxygenated blood to the left ventricle, that in turn pushes it to the body through the Aorta.

This de/oxygenated blood is pushed between the heart champers by the contraction of the heart
This heart wall, as any other muscle, is composed of muscle cells, which is aligned in a network. These muscle cells are divided into two types: the myocardial contractile cells, and the myocardial conducting cells. The former one conducts the electrical pulses in the heart and is responsible for contractions. Whereas the later one forms the conduction system of the heart [1].

![Human heart conduction system](image)

Figure 2.1: Human heart conduction system [1]

The conduction system of the heart has five main components: the sinoatrial node (SA), the atrioventricular node (VA), the atrioventricular bundle (bundle of His), the atrioventricular bundle branches, and the Purkinje cells as illustrated in figure 2.1. The electrical activity of this conduction system is as follow:

1. The conduction system is at the rest state.

2. The SA node is composed of a mass of myocardial conducting cells that are responsible for the normal cardiac rhythm creation. It is also known as the pacemaker of the heart since it initiates the electrical pattern or sinus rhythm that is followed by the contraction of the heart [1].

3. The impulse travels from the SA node to the left atrium through the Bachmann’s bundle, and reaches the AV node.

4. The impulse reaches the AV node and experience a delay to make sure that the blood was pumped effectively from the atria to the ventricles.

5. After that delay, the impulse travels through the bundle branches to the Purkinje fibers. Then, it reaches the right papillary muscle through the moderator band. The impulse reaches the
left papillary muscle approximately at the same time. However, there is no corresponding moderator band in the left ventricle [1].

6. The impulse then reaches the apex and spreads to the contractile fibers of the ventricles.

7. A contraction of the ventricles begin to pump the blood to the body or the lung to be oxygenated.

By using the proper electrodes, these electrical activities can be recorded. This tracing of the activity is the electrocardiogram (ECG) that we usually see in the clinic and what this thesis depends on. Figure 2.2 illustrates the normal ECG waveform.

![Figure 2.2: Cardiac heartbeat trace and wave points.](image)

There are five main fiducial or characteristic points in the ECG beat: P point, Q point, R point, S point (together forms the QRS complex), and T point. Each wave segment corresponds to an event in the cardiac cycle as depicted in figure 2.3. The P wave is of a small amplitude and represents the depolarization of the atria. The atria begin to contract after 25 milliseconds of the start of the P wave. The QRS complex represents the depolarization of the ventricles. It has a large signal amplitude since the ventricles have large muscle size compared to the atria. The ventricles start to contract after the R peak. The repolarization of the atria happens during this wave; however, it is masked by the ventricular depolarization. Finally, the T wave which represents the repolarization of the ventricles.

However, the adjustment of the ECG electrodes placed on the body can affect the recorder signal. For instance, the improper localization of the electrodes may affect the amplitude of the measured electricity, as well, the impedance between the unadjusted electrode and the body surface.
may influence the measured signal. Therefore, a standardization of the ECG leads is an essential task for recording the individual’s ECG.

### 2.2 Recording the ECG signal

As mentioned earlier, the electrical activity of the heart can be measured by placing some electrodes on the body surface of the individual. These electrodes are placed to capture the spatio-temporal changes during the cardiac cycle [10]. The term ‘electrode’ refers to the conductive pad in contact with the body (i.e. the sensor) that we place on the body surface, while the term ‘lead’ refers to the voltage difference between two electrodes.

In order to globalize the understanding of the ECG signal, a standardization should be set for the usage and positioning of those electrodes. Different ECG lead configurations have been investigated. For instance, the orthogonal lead system, the standard 1-, 3-, 6-, and 12-lead configurations.

The choice of the lead configuration system not only depend on the extracted information, but practical utilization is also considered. For instance, in some systems, it is hard to connect 10 electrodes to the individual, such as biometric systems, on the other hand in the diagnostic ECG, we can use the 12-lead system.

Before proceeding with the lead configurations, let us define the positions of the ECG electrodes and their acronyms as in Table 2.1

1. **12-lead ECG System**: The 12-lead ECG system is the most common configuration in clinical
Table 2.1: ECG electrodes acronyms and positions.

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Electrode Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA</td>
<td>Right arm</td>
</tr>
<tr>
<td>LA</td>
<td>Left arm</td>
</tr>
<tr>
<td>RL</td>
<td>Right leg</td>
</tr>
<tr>
<td>LL</td>
<td>Left leg</td>
</tr>
<tr>
<td>V1</td>
<td>Fourth intercostal space to the right of the sternum</td>
</tr>
<tr>
<td>V2</td>
<td>Fourth intercostal space to the left of the sternum</td>
</tr>
<tr>
<td>V3</td>
<td>Between leads V2 and V4</td>
</tr>
<tr>
<td>V4</td>
<td>Fifth intercostal space</td>
</tr>
<tr>
<td>V5</td>
<td>Left anterior axillary line, and horizontally even with V4</td>
</tr>
<tr>
<td>V6</td>
<td>Midaxillary line, horizontally even with V4 and V5</td>
</tr>
</tbody>
</table>

Table 2.1 details the acronyms and positions of ECG electrodes. The table is as follows:

- RA: Right arm
- LA: Left arm
- RL: Right leg
- LL: Left leg
- V1: Fourth intercostal space to the right of the sternum
- V2: Fourth intercostal space to the left of the sternum
- V3: Between leads V2 and V4
- V4: Fifth intercostal space
- V5: Left anterior axillary line, and horizontally even with V4
- V6: Midaxillary line, horizontally even with V4 and V5

routines [10]. It uses only ten electrodes. This 12-lead system consists of a combination of three lead configurations:

(a) Three bipolar limb leads: These leads are denoted as lead I, II, and III. They are mainly measured using the electrodes on the right arm (RA), left arm (LA), and left leg (LL) as shown in figure 2.4, where the voltage is measured as follows:

\[ I = V_{LA} - V_{RA} \]  \hspace{1cm} (2.1)
\[ I = V_{LL} - V_{RA} \]  \hspace{1cm} (2.2)
\[ I = V_{LL} - V_{LA} \]  \hspace{1cm} (2.3)

This is a common placement of the electrodes that is based on the Einthoven’s triangle, an imaginary triangle drawn around the heart. Einthoven’s law states that if the values for any two points of the triangle are known, the third can be estimated.

\[ II - I = III \]  \hspace{1cm} (2.4)

Therefore, sometimes it is not necessary to measure lead III, since we can induce it from the other bipolar limb leads.

(b) The augmented unipolar limb leads: These leads are denoted as aVF, aVL, and aVR.
They use the same electrodes as the bipolar limb leads, but the voltage difference is between one of the electrodes and the average of the other two as shown in figure 2.4. These augmented leads capture the ECG signal which is shifted 30 deg from the bipolar limb leads. They are considered as unipolar leads, since we measure one electrode while averaging the other two. Similarly, we can estimate the values of these leads from leads I, and II.

![Image](image_url)

Figure 2.5: The precordial leads in ECG recording system.

(c) The precordial leads: The six precordial leads, named V1 to V6, are placed on the front and left side of the chest. They lie in the transverse (horizontal) plane, that is perpendicular to the other six leads. The six precordial electrodes act as the positive poles for these six precordial leads, while the Wilson’s central terminal $V_{WCT}$ is used as the negative pole as in figure 2.5. Leads V1 and V2 represent the activity of the right ventricle. Whereas leads V3 and V4 the activity of the frontal part of the left ventricle, while the leads V5 and V6 represents the lateral part of the left ventricle.

With this lead configuration, all the heart activity can be measured since all the directions were covered. Therefore, more assessment of the rhythmic and ischemic cardiac abnormalities can be provided [40]. Figure 2.6 shows the shape of the heartbeat using the different leads.

2. **6-lead ECG System:** In this lead configuration, we use the three bipolar limb leads in addition to the three augmented limb leads. Therefore, leads I, II, III, aVF, aVL, and aVR. As mentioned previously, we can derive the augmented leads from the bipolar limb leads. Hence, we can use only the three primary electrodes.
3. **3-lead ECG System:** In this system configuration, lead I and lead II are measured, while lead III is derived as in equation 2.4. Therefore, only the bipolar limb leads are used, or three electrodes.

4. **1-lead ECG System:** This simple lead configuration is more practical than other configurations, since it needs only two electrodes. It can also be used to record regular exercise, cardiac events, resting cardiac state, and sports activities. The advantage of using this lead configuration besides being practical is that it can be found at low price and that it became available for personal use [41]. Leads I is the most commonly collected vector, as it is measured using the reference of the left arm (LA) to the right arm (RA) as mentioned before. It provides the basic heart monitoring and measure the potential difference between the two arms. In the ECG recording of that lead, a wave a depolarization moving in the direction of the left arm gives a positive deflection since the positive electrode is on the left arm. Moreover, the maximal positive ECG deflection occurs when a wave of depolarization travels parallel to the axis of the lead (i.e. between the right and left arms). Therefore, most of the cardiac diseases can be interpreted from this lead, although other leads are necessary in some cases where confirmation and another perspective are needed.

This configuration is more practical in some application, where placing many electrodes on the individual is not possible.

5. **Orthogonal Lead System:** This lead system reflects the heart activity in the three main perpendicular directions (i.e. X, Y, and Z). Additional information can be gained by visualizing the electrical activity in a three-dimensional view along with its projection onto the XY-, YZ-, and XZ-plane.
2.3 ECG Abnormalities

By choosing the appropriate ECG lead configurations, physicians can diagnose the ECG waveform and interpret visually the individual’s health condition. However, this needs a training and full understanding of the ECG complexities and intricacies, which usually requires experience in it.

For instance, as we discussed before that the electrical activity starts from the SA node in the right atrium, in some cases, another area in the heart initiates the impulse rather than the SA node. This area is usually called ectopic pacemaker or ectopic focus [1]. This ectopic focus may be developed due to the consumption of certain drugs, or any other pathological conditions. Occasional occurrence of this ectopic focus is not a life threatening. However, chronic ectopic focus may lead to arrhythmia, an abnormal heart pattern that is due to uncoordinated heartbeats [1].

Other diseases such as the left bundle branch block (LBBB) can also be found, where a blockage is found in the electrical impulse pathway bundles to the left ventricle. This blockage may cause a delay in the ventricle contraction. Therefore, a deficiency in blood pumping to the body and may lead to a heart attacks, high blood pressure, viral bacterial infection in the heart muscle, or complete blockage of the electric impulse [42]. This blockage can also be found in the right bundle pathway that triggers the right ventricle. In that case, the disease is called right bundle branch block (RBBB). The LBBB can be seen in the ECG record as a signal with wide QRS complex (i.e. ventricles contraction) due to the aforementioned delay, or a notched R wave (i.e. appears as 'M'-shaped) as the two ventricles depolarize at different times. Figure 2.7 shows how we can interpret these abnormalities from different ECG leads.

Another diagnosis that can be interpreted from the ECG recording is the heart rate (HR), which is frequency of the QRS complexes, and it is commonly referred to as the "heart rate".

The normal heart rate is between 60 and 100 beats per minute (bpm). Bradycardia is when the heart rate drops below 60 bpm. Whereas tachycardia is the condition in which the heart rate goes above 100 bpm in the resting state.

The bradycardia, low heart rate, may be caused by either inherent factors such as older individuals, or external causes to the heart such as pathological conditions of thyroid, neurological disorders, or autoimmune pathologies. On the other hand, tachycardia, high heart rate, may be normal in pregnant women and individuals that experience stress [1].
Figure 2.7: Interpretation of left and right bundle branch block diseases from ECG signals of different leads. [3].

2.4 Summary

In this chapter, we explored the anatomy of the heart and its chambers. The cardiac cycle was represented at its corresponding fiducial point in the ECG signal. A brief discussion about different ECG lead configuration was given. Finally, some cardiac abnormality that might be found in the ECG are studied.
Chapter 3

Literature Review

For several decades ECG has been utilized as a diagnosis tool for the cardiac health status information. Nevertheless, some studies revealed the fact that these signals can be used in the biometric system. Not only having a high accuracy but it provides some advantages over other biometric traits.

In this chapter we will discuss the literature review proposed in both arrhythmia detection as well as ECG-based biometric systems.

3.1 ECG in Diagnosis

There are several computer-aided approaches in the literature that extract a set of features from an original ECG signal to detect arrhythmia. These approaches can be categorized into three main groups based on the type of features used for processing by the classification algorithm: time-domain, frequency-domain, and time-frequency domain features.

3.1.1 Time Domain Features

The time-domain features employ the amplitude of the ECG signal that changes with time. For example, Mazhar et al. [43] classified 15 different cardiac heartbeat types using the statistical features of the temporal ECG signal. First, they extracted the QRS fiducial points from the signal, and then statistical features such as the mean, standard deviation, and the energy are used to feed a support vector machine classifier. The classification is done one two stages. The first stage is to separate normal and abnormal beats. Whereas the second stage is to classify the abnormal heartbeats into different arrhythmia classes. Their algorithm achieved an accuracy of 98.78% for 100 iterations.
Dewangan, and Shukla [44] proposed an algorithm where they first detected the P-R-T fiducial points using a windowing technique. Then, morphological features were extracted including the R peak amplitude, PR interval, and the wavelet transform coefficients to feed a neural network classifier.

In addition, Yu SN, and Chou KT [45] proposed a method to classify ECG signals using independent component analysis (ICA) coefficients and RR intervals as features, and a neural network as a classifier.

Pan et al. [6] used a Hidden Markov Models (HMM) to classify three different types of heartbeats, such as bundle branch block, ventricular premature contraction, and atrial premature contraction using MIT-BIH database. They compared the performance of their HMM-based classifier to other ones (such as neural networks, and restricted coulomb energy networks). Although their proposed approach outperformed other methodologies, they combined both the left and right block bundle into one group.

Chazal et al. [46] proposed a method to classify five different ECG heartbeats types. The cohort involved 44 subjects. Morphological features including RR interval, QRS onset and offset, T-wave duration of the time signal were utilized. The algorithm achieved an accuracy 81.9%.

Yu and Chou [45] proposed an algorithm to classify eight different heartbeat types. They utilized the ICA coefficients, specifically 23 coefficients, as well as the RR interval as features to feed a three-layer neural network classifier. The accuracy reported for the proposed system was 98.37% based on the MITDB arrhythmia database.

Some researchers have utilized the deep learning to build a patient-specific model to detect arrhythmia diseases. For instance, Zhang et al. [47] used the recurrent neural network (RNN) and density-based clustering techniques to propose a patient-specific ECG classification system. They utilized the RNN to explore the correlation between the ECG time signal points. Clustering is then used to find the representative beats for the training of patient-specific model. An average accuracy of 99.7% was achieved in the classification four classes.

Cheng and Dong [48] proposed an approach that utilized the personalized features for each patient. The ECG is first preprocessed. Then, the correlation coefficients are extracted from each heartbeat. Six statistical features related to the correlation coefficients are extracted including the average, median, minimum, maximum, and standard deviation, as well as the statistical features of the RR intervals of the heartbeats in the patient’s ECG signal. An average accuracy of 95.46% is achieved on three databases.

Moreover, Rajpurkar et al. [49] used a deep convolutional neural network to classify sinus rhythm
and atrial fibrillation from 30-second ECG input segments. The output was compared to a cardiologist committee. A dataset of 30000 patients was collected by the authors and the aggregate F-score of the proposed algorithm was 80.9% when compared to the real physicians.

3.1.2 Frequency Domain Features

In the frequency-based features, the ECG signals are converted from the time-domain into the frequency domain, then the features are extracted from the latter such as the Fourier transform [50, 51]. For instance, Desai et al. [52] proposed such a method that first, processes the discrete cosine transform coefficients extracted from the ECG signal, followed by independent component analysis to reduce the large number of features. Their method achieved 95.98% accuracy on the MITDB dataset.

Minami et al. [53] studied the change in the QRS complex of the different arrhythmia types. The Fourier transform coefficients were extracted from the QRS complex and then fed to a neural network for classification. A set of 124 records of lead I was used in the algorithm validation. An accuracy of 98% was achieved in their approach.

Gothwal et al. [54] also utilized the fast Fourier transform to extract the features from 40 records of the MITDB database. The Fourier transform was mainly used to extract the fiducial points of the ECG signal such as Q, R, and S points. Then features such as RR interval (interval between R peaks), and QR duration are estimated from the extracted wave points.

3.1.3 Time-frequency Domain Features

In this type of features, both the time and frequency domains are studied. It can be seen as investigating the changes in the signal frequencies along the time. The discrete wavelets transform (DWT), continuous wavelets transform (CWT) [55, 56], and discrete cosine transform (DCT) [57] are examples of transformations that allow us to study the signal in the time-frequency domain. For instance, Can Ye et al. [58] proposed an arrhythmia detection approach using the dynamic features of the ECG signal. The signal is first preprocessed to remove the noise. Then, the wavelet transform coefficients are extracted followed by applying ICA to reduce the number of coefficients extracted. The RR intervals’ features are then combined with the wavelets coefficients to feed a support vector machine (SVM) classifier. The average accuracy achieved reached 99.6%.

Thomas et al. [59] employed the dual tree complex wavelet transform (DTCWT) to extract features from the ECG signals to detect cardiac arrhythmia. The features from the fourth and
fifth scale of the tree of the SQR complex were used along with other time information such as the skewness and kurtosis or the heartbeat signal. A classifier of neural network was used in the abnormal heartbeat detection. An overall sensitivity of 94.64% was reported using the DTCWT compared to 91.23% using the discrete wavelet transform.

3.2 ECG in Biometrics

The literature work in the ECG biometrics can be divided into three main groups: the fiducial-based systems, non-fiducial-based systems, and hybrid systems.

3.2.1 Fiducial Features

In this category, the features are extracted by utilizing the detection of all or part of the characteristic points of the temporal ECG. These features may include the amplitude, area, angle, or dynamics between the heartbeats.

These characteristic points are when the heart electrical events take place. For instance, the R characteristic point that represents the depolarization of the left ventricle can be used to extract the RR interval feature.

For instance, Biel at el. [60] utilized the 12 ECG leads to identify 20 subjects. They measured the ECG activity for each subject for at least 5 times to guarantee that the positioning of the electrodes does not affect their system’s performance. Thirteen fiducial features including the P and QRS wave onsets, P, Q, R and S duration were extracted from each wave. Then, they employed the correlation matrix to reduce the number of features to only 12 features. A soft independent modeling of class analogy (SIMCA) classifier is then used to identify the subjects using the selected features. An average correct classification rate of 98% was achieved using different number of selected features.

Kyoso and Uchiyama [61] proposed an approach to identify nine subjects from one the limb ECG lead II. They extracted the characteristic points of the ECG signal using the second order derivative of the signal. Fiducial features such as the P duration, PQ interval, QRS interval, and QT interval were extracted. Then different combination of pairs of features are then used in the discrimination analysis using Mahalanobis’ generalized distance. An average accuracy of 96.46% was achieved for eight subjects where the ninth subject obtained 76.1% when using the combination of the QRS-QT interval features.

Following their work, Kim et al. [62] utilized the QRS, QT, and RT intervals of the time-domain ECG to identify ten normal subjects using Mahalanobis’ distance. Their system showed the
feasibility to use the ECG to identify individuals.

Irvine et al. [63] proposed a biometric system based on the ECG signal to study the effect of the emotional and mental state variation on the identification performance. They carried out two experiments. Each experiment had different task that represent different mental or emotional state. Signal preprocessing to remove the high and low frequency noises were performed. Then, extraction of nine fiducial points from the temporal ECG signal took place. Different fiducial features are then defined including the L,P',S'T', and QT intervals yielding a total of fifteen feature. A stepwise discriminant analysis was done to reduce the number of features. Afterwards, a linear discriminant classifier was used in the identification process to test inter- and intra-task performance. Their results showed that ECG can be used in the biometric process and that the features which they utilized might need a non-linear classifier instead of the linear discriminant one.

Using the same dataset, Israel et al. [64] employed the ECG signal to identify individuals. The signal was first preprocessed using bandpass filter. Then, the fiducial points were detected in a beat-by-beat basis. Fifteen features were then extracted with respect to the R peak. These features include the RQ, RS, RT intervals, as well as the P,T waves durations. A stepwise canonical correlation was then used to reduce the number of features that were fed to a linear discriminant classifier. They concluded that the emotional state change of the individual which is reflected on the ECG can be overcame by normalization of the features. Their algorithm achieved an 82% and 79% heartbeat identification rate using different lead locations, and average accuracies of 80.1% and 64.5% using different anxiety state. They also used the same dataset to study the minimum number of heartbeats needed to verify individuals [65]. The ECG characteristic points were identified, and fiducial features are then extracted. These features include the fiducial points' amplitudes and the RR, PQ, and QT intervals. A sequential procedure was then employed to estimate the minimum number of beats required to verify each subject using different accepted error rate. The calculated minimum heartbeat at 1% allowable error was 2 with a mean of 3.38 and maximum of 8.

Zhang and Wei [66] on the other hand, used the Bayes classifier to identify 502 subjects. The fiducial features included the ECG characteristic points amplitude and duration as well as fiducial intervals such as QRS and PR intervals. Different experiments were carried out on four ECG leads including Lead I, II, V1, and V2. Whereas Lead V2 obtained the highest accuracy of 79%, and lead I achieved 85.3%.

Using neural networks, Palaniappan and Krishnan [67] identified ten subjects using the MITDB Normal Sinus database. The ECG signal was preprocessed to remove the noise. Then, the fiducial points were detected to extract the features. Six features were identified involving the RR, QRS
intervals as well as QRS form factor [68] to feed a multilayer perceptron (MLP).

Guennoun et al. [69] used the ECG signal to verify if the specified user is the actual user during a whole session of time. A set of 16 subjects were involved in the study with recordings of 15 minutes each. The ECG signal is first preprocessed to remove the noise. Then, the fiducial points were segmented. Afterwards, twenty-four features were extracted including fiducial intervals and amplitudes, followed by a feature reduction to produce only nine features. Mahalanobis’ distance classifier is then used to determine whether the current user is the actual user during the specified session.

Venketsh et al. [70] on the other hand, used the ECG fiducial features to identify 15 subjects. The QRS wave was detected using the Pan and Tompkins algorithms [71] followed by detection of the P, T, Q, and S points using the zero-crossing of the ECG signal. Then, a set of fiducial features include the PR, QRS, and ST intervals were extracted. Different classifiers were used for identification including one-stage and two-stage classifiers that involved dynamic time warping and Fisher’s linear discriminant with nearest-neighbor classifier. An identification accuracy of $94 \pm 2\%$ was achieved.

Shen et al. [72] used the ECG signal of 20 subjects for verification purpose. The signal was first preprocessed, then the fiducial points were extracted. A set of seven fiducial features were derived from the signal including the T amplitude, QT and QR intervals. Afterwards, a template-based procedure was performed to compare 20 beat per subject with the subject’s template beat. This yielded an accuracy of 95%. The authors provided an alternative approach by integrating decision-based neural networks to improve the accuracy. The average value of cross-correlation coefficients was measured between the template and the 20 beats. If this value is less than a threshold (i.e. 0.85), then a non-match decision is taken. On the other hand, if the average value is above this threshold, then they use the neural network for decision making. This integration increased the accuracy to 100%.

### 3.2.2 Non-fiducial Features

In this category there is no need to use the fiducial points of the ECG to extract the features. Instead, one or two fiducial points only can be used for heartbeat segmentation. The extracted features are usually obtained from the spectral domain (i.e. wavelet transform, discrete cosine transform). This sometimes yields a better approach since the P and T characteristic points are affected by the noise and the health condition of the individual. For instance, Bassiouni et al. [73] studied the
identification of 30 subjects using the discrete cosine transform (DCT). The DCT coefficients of the autocorrelation of an ECG signals of 10 seconds were extracted. Then the first 400 coefficients were used to feed a neural network classifier. An accuracy of 96.7% was achieved the utilized scheme.

On the other hand, Camara et al. [74] used a dataset of 18 subjects’ ECG signals for identification. The Hadamard transform [57] was applied to the two leads’ ECG signal, where only 24 coefficients were used per lead. This Hadamard transform has the same idea of projecting the signal on a set of orthogonal sinusoidal waves, but instead, it uses square waves called Walsh functions. They studied the performance of these 48 features alone and with the incorporation of the Shannon and Log-energy entropy. A K-nearest neighborhood classifier was used in the classification process for simplicity, where the average accuracy obtained for the 48 coefficient features was 94.2% and 96.6% by incorporating the two entropies.

Chan et al. [75] collected a three-session ECG data of 50 subjects for identification. They used the thumbs ECG leads instead of the chest leads. The signal was preprocessed to remove the noise using a notch filter. The fiducial points, PQRST points, were then detected. A temporally alignment was then performed on the signal using the cross-correlation coefficients, followed by estimation of the signal-averaged ECG (SAECG). Three measures were then used for the similarity. These measures include the percent residual different, correlation coefficient, and wavelet distance measure. Using these measures, an accuracy of 89% was obtained by testing on different session.

Hejazi et al. [76] studied the verification of individuals using ECG signals. The cohort involved 52 healthy subjects, where all the subjects contributed with 3 minutes ECG recording while only 16 had additional sessions. This resulted in a total number of 68 recordings. The ECG signal is first preprocessed to remove the noise using discrete wavelet transform and hard threshold rule. Then, the signals from each subject is then segmented into non-overlapped windows of five seconds each. Afterwards, the normalized autocorrelation coefficients are then extracted from each window that account for 1028 coefficients. Linear and non-linear feature reduction algorithms were used along with one-class SVM classifier.

Fatemian and Hatzinakos [77] used two public ECG databases to perform individual identification. First, the ECG signal is resampled to 250 Hz, and wavelet reconstruction at the 3rd scale followed by a moving average smoothing was done in order to remove the noise. Then, the detection of the R peak is performed followed by detection of Q,R,S, and T wave points. The ECG heartbeats of each subject is then aligned, and the median of multiple beats was taken as the gallery for that subject. An accuracy of 99.62% was achieved on the tested datasets.

Chan et al. [78] used the chaotic property of the ECG to identify individuals. A set of 9 subjects
were used in the identification process. Six different features were used in the system including four Lyapunov exponents, correlation dimension, and root mean square level of the ECG signal. The outliers were then removed from the dataset, and the features are then fed to a neural network. An identification accuracy of about 90% were achieved using 90%, 10% training and testing partitioning.

Chiu et al. [79] utilized on ECG lead to perform identity verification. Two datasets were used: one that contains 35 normal subject and 10 arrhythmic subjects, while the second one is measured using their own measured ECG signals that consisted of 25 subjects. The ECG signal is first preprocessed using a low-pass filter and high-pass filter of 0.5 and 45 Hz cutoff frequencies. Then, the detection of the QRS complex took place using So and Chan method [80]. After the detection of the R peaks, a set of 43 and 84 signal points were taken before and after the detected peak to represent a beat. Signal synthesis then is done by concatenating four beats, yielding a signal of 512 data point. Discrete wavelet transform is then applied to this 512-point signal using level-9 decomposition. The resulted vector of coefficients is used to compare the subject in question with the database using the Euclidean distance. An accuracy of 100% was achieved on normal subjects, while the accuracy of the arrhythmic subjects was 81%.

Agraffioti and Hatzinakos [37] studied the ECG irregularities in the identification and authentication of individuals. The ECG signals from three public databases were preprocessed and segmented into windows. Autocorrelation is then performed on each window. The normalized autocorrelation coefficients are then subjected to complexity measure to detect the premature ventricular contraction ECG segments and eliminate it from the classification process. Linear discriminant analysis was then done to reduce the number of features, the autocorrelation coefficients. Then, K-nearest neighbor classifier was used for individual verification and identification. They later utilized the same algorithm to study the effect of integrating 12 leads information [81]. The study cohort involved 14 healthy subjects. Each subject had two recordings of ECG. The ECG inputs of 5 seconds window were first filtered to remove the noise. Then, the autocorrelation coefficients were extracted from the window. Afterwards, a linear discriminant analysis was performed to reduce the number of features (i.e. autocorrelation coefficients). A K-nearest neighbor classifier was then used to assign the input recording to the matching individual using Euclidean distance. The accuracy was estimated for those subjects that the classifiers agreed to assign to certain individual. Having this accuracy estimation, an identification rate of 100% was achieved while having 21.43% subject rejection rate.

Chen et al. [78] used their developed portable instrument to measure their own ECG dataset for human identification. The dataset involved 9 subjects in rest state. A set of six features were used in the identification process. These features include four Lyapunov exponents [82], root mean square
level (RMS) to study the magnitude of the varying ECG signal, and the correlation dimension [83].
An identification accuracy of 90% was achieved using neural networks as a classifier.

Hejazi et al. [84] proposed an ECG-based biometric system that utilizes the autocorrelation along with linear dimension reduction. The cohort of the study involved 52 subjects. The signal is first denoised using discrete wavelet transform (DWT). Several feature reduction methods including linear discriminant analysis and principle component analysis (PCA), as well as kernel principle component analysis (KPCA) were tested. The reduced features are then used to feed an SVM classifier to identify the subjects. The results showed that using the Gaussian KPCA as a feature reduction method for the autocorrelation coefficients achieved the lowest false non-matching rate of 4.83 in case of one session and 2.297 in case of two sessions.

Poree et al. [85] studied the identification and verification of individuals using the ECG signals. They utilized a 12-lead configuration system. The signals are first preprocessed, then the R peaks are identified. Afterwards, the average of 10 heartbeats are taken to estimate the heartbeat shape. The scores are determined by computing of the correlation coefficients between the pairs of shapes of the ECG signals.

Nevertheless, deep learning has gained attention of the researchers in the last couple of years. Zhang et al. [86] used 1D convolutional neural network to identify subjects using different combination of wavelets at different scales. The algorithm was tested on different ECG databases and an average identification rate of 93.5% was achieved all over the databases.

3.2.3 Hybrid Features

Dar et al. [87] used different databases to utilize the ECG biometric in identification. The ECG signal is first preprocessed by using a 6th order polynomial curve fitting followed by detection of the R peaks. A training set that consists of 35 seconds for each subject were taken to extract the features. These hybrid features include both heart rate variability features and discrete Haar wavelet. The heart rate variability features include the RR interval mean, RR interval standard deviation, heart rate standard deviation, etc. Those features are then fed to a best search algorithm [88] to select the uncorrelated features. This resulted in a set of 46 features. A classifier of random forest is then used to identify the individuals. System accuracies of 95.85%, 100%, and 83.88% were achieved using MIT-BIT arrhythmia, MIT-BIH normal sinus, and ECG-ID databases respectively.

Wang et al. [89] utilized both the fiducial and non-fiducial features in their ECG identification system. The fiducial features included the P,Q,R,S and T position and amplitude. Whereas the
non-fiducial features employed the autocorrelation coefficients and discrete cosine transform.

3.3 Summary

In this chapter, we went through the literature work of both clinical application and biometric application of ECG. The clinical work can be categorized into: time-domain, frequency-domain, and time-frequency-domain. Whereas the biometric-based is divided into fiducial-based, non-fiducial-based, and hybrid approaches.
Chapter 4

Methodology

4.1 ECG Databases

Many ECG databases have been proposed in the literature using different lead configurations. These databases can be classified into two main categories: public and private databases. The public databases were made available for the researches to access. Whereas private databases were acquired by the researches for their own researches. In both categories, the protocol of the acquisition depends mainly on the purpose of the ECG signal. For instance, databases that use 12-lead configuration are mainly designated for the purpose of diagnosis. Whereas, most of the biometric ECG applications require one- or two-lead configuration for practical reasons. However, all the databases that we consider here were obtained using in vitro acquisition devices where the electrodes are attached to the surface of the body during the recording process. Table 4.1 shows some of the ECG databases in the literature work.

Most of the ECG databases made public was obtained for clinical use. Therefore, they do not have multiple-session protocol, where the patients’ readings have to be done in different times. These databases are good for diagnosis. However, in the biometric a more intra-subject variability can be further studied if more than one session we carried out.

MIT-BIH Arrhythmia Database (MITDB)

MIT-BIH arrhythmia database (MITDB) [90] is an ECG database that was recorded in the laboratories at Boston’s Beth Israel Hospital (now the Beth Israel Deaconess Medical Center) and at MIT using a battery-based portable Holter monitor that records the heart’s electrical activity.
Table 4.1: ECG databases in the literature work.

<table>
<thead>
<tr>
<th>Database</th>
<th>Type</th>
<th>Subjects</th>
<th>Condition</th>
<th>Sessions</th>
<th>Frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MITDB [90]</td>
<td>Public</td>
<td>47</td>
<td>N/AbN</td>
<td>1</td>
<td>360</td>
</tr>
<tr>
<td>CEBSDB [91]</td>
<td>Public</td>
<td>20</td>
<td>N</td>
<td>1</td>
<td>5000</td>
</tr>
<tr>
<td>NSRDB [92]</td>
<td>Public</td>
<td>18</td>
<td>N</td>
<td>1</td>
<td>128</td>
</tr>
<tr>
<td>STDB [93]</td>
<td>Public</td>
<td>28</td>
<td>N/AbN</td>
<td>1</td>
<td>360</td>
</tr>
<tr>
<td>Fantasia [94]</td>
<td>Public</td>
<td>40</td>
<td>N</td>
<td>1</td>
<td>250</td>
</tr>
<tr>
<td>AFDB [95]</td>
<td>Public</td>
<td>23</td>
<td>N</td>
<td>1</td>
<td>250</td>
</tr>
<tr>
<td>VFDB [96]</td>
<td>Public</td>
<td>22</td>
<td>N/AbN</td>
<td>1</td>
<td>250</td>
</tr>
<tr>
<td>PTBDB [97]</td>
<td>Public</td>
<td>290</td>
<td>N</td>
<td>&lt;5</td>
<td>1000</td>
</tr>
<tr>
<td>UofTDB [98]</td>
<td>Private</td>
<td>1020</td>
<td>—</td>
<td>&lt;5</td>
<td>200</td>
</tr>
<tr>
<td>Zhang and Wei [66]</td>
<td>Private</td>
<td>520</td>
<td>—</td>
<td>1</td>
<td>500</td>
</tr>
<tr>
<td>Shen et al. [99]</td>
<td>Private</td>
<td>168</td>
<td>—</td>
<td>1</td>
<td>500</td>
</tr>
<tr>
<td>Chan et al. [75]</td>
<td>Private</td>
<td>50</td>
<td>—</td>
<td>3</td>
<td>1000</td>
</tr>
</tbody>
</table>

N=Normal, AbN=Abnormal heart condition.

Continuously.

MITDB contains 48 ECG recordings of 47 subjects. Each recording is half-hour of two-channel ambulatory ECG that was sampled at 360 Hz. As for the annotation, two or more cardiologists independently annotated each record.

**CESBS Database**

The purpose of the combined measurement of ECG, breathing and seismocardiograms database (CEBSDB) [91, 100] was to check the effect of breathing on the detection of the RR time series, and to compare the detected RR time series of the ECG signal to the seismocardiogram.

The cohort involves 20 healthy subjects, and the ECG records were obtained in supine position using the Biopac MP36 data acquisition system. The experiment involved three phases. First, the ECG recording was performed in a still position for 5 minutes. Second, the ECG recordings were taken while the subjects are listening to classical music for 50 minutes. Finally, after the music has ended another 5 minutes ECG recordings were obtained. The obtained ECG signals are two channels (i.e. lead I and II) in the recording sampled at 5 KHz.

**MIT Normal Sinus Rhythm (NSRDB)**

MIT Normal Sinus Rhythm (NSRDB) [92] includes the ECG recordings of 18 subjects, five men and thirteen women, obtained the Laboratory at Boston’s Beth Israel Hospital (now the Beth Israel Deaconess Medical Center). The subjects included in this database do not exhibit significant arrhythmias. The ECG signals obtained were sampled at a frequency of 128 Hz on two channels (lead I and II).
MIT-BIH ST Change Database (STDB)

The MIT-BIH ST Change Database (STDB) [93] involves the variable length ECG recording of 28 subjects. Most of the recordings were obtained during exercising stress tests that exhibit transient ST depression. Five records of the 28 exhibits ST elevation (the ST segment is abnormally higher than the base line). The recordings were obtained from two channels that were sampled at 360 Hz.

Fantasia Database (Fantasia)

The Fantasia database [94] was obtained from subjects while watching the Disney movie Fantasia (120 minutes). Forty subjects including 20 young (21 to 34 years old) and 20 elderly (68 to 85 years old) were involved in the database measurements, where each group has the same number of men and women. The acquired ECG signal is of one channel at sampling frequency of 250 Hz.

MIT Atrial Fibrillation Database (AFDB)

The atrial fibrillation database (AFDB) [95] was obtained at Boston’s Beth Israel Hospital using ambulatory ECG recorders. It contains 25 ECG recordings of 23 subjects. Each recording is of 10-hour duration and contains two ECG channels sampled at 250 Hz.

MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB)

The ventricular arrhythmia database (VFDB) [96] was obtained to recognize ventricular tachycardia, flutter, and fibrillation from electrode motion noise (i.e. four classes). It includes half-hour ECG recordings of 22 subjects sampled at 250 Hz. Therefore, this database can be used to validate the system in case of heart abnormalities since the QRS complex is hard to extract in case of the diseases presented in this database.

PTB Diagnostic ECG Database (PTDB)

This database was obtained by the Physikalisch-Technische Bundesanstalt (PTB), National Metrology Institute of Germany [97]. It was collected from healthy subjects as well as patients with different heart diseases.

The database contains 549 records with different lengths obtained from 290 subjects sampled at 1 KHz. The age of the subjects involved ranges from 17 to 87, with a mean of 57.2 for men and 61.6 for women. Each subject has one to five records, where each ECG record has 12-lead configurations in addition to three Frank leads.
Among the 290 subjects, 148 subjects suffer from myocardial infarction, 18 have cardiomyopathy or heart failure, and 52 healthy subjects.

The challenge in this database is the number of subjects. For instance, unlike the other databases that usually have at most 50 subjects, PTBDB has 290 individuals, which mimics real-world scenarios.

4.2 ECG Signal Processing

One of the challenges with the biological signals, specifically ECG signals, is that it is contaminated with noise. During the early work of ECG signal and computer processing, it was evident that the interpretation of the signal depended on the accuracy of the signal measurement and its noise effect [10, 101].

The most three common noise sources in the ECG signals are the baseline wander, the power-line interference, and the muscle and respiration noise. However, due to the variation of these noise factors during the acquisition, it is hard to provide a universal ECG signal analysis algorithm that works for all the acquisition conditions [10]. Therefore, the complexity of each processing algorithm as well as the sequence of its steps depends on the application. However, some concerns should be taken into account when the noise is removed from the signal. For instance, the noise removal step should maintain the desired information undistorted, which is another challenge. Another concern is the effect on the QRS complex, since the QRS is considered as a peak signal with a large amplitude, the filter design should be aware that this complex is part of the original signal and not a noise.

4.2.1 Noise Removal

The noise removal step is an important step before the investigation the ECG signal. Figure 4.1 shows the noisy and filtered ECG signal using the moving means filter.

Baseline Wander

Several reasons can contribute in the baseline wander noise such as the cables moving during recording, patient movement, or loose electrodes. The frequency of the baseline wander is usually below 0.5 Hz [10]. However, this frequency may increase if the patient motion increased for instance during stress tests.
Powerline Interference

This type of noise is due to the alternating current in the surrounding environment. The usual frequency for this noise is either 60 or 50 Hz. A linear filtering using the notch filter \([102, 103]\) can be applied to the ECG signal to remove this effect of the powerline.

Muscle Motion Noise

This type of noise is very common in the ECG signal especially in the stress tests. It comes from the fact that the recorded signal of the ECG might be affected by the electrical activity of the surrounding muscles. Unlike the powerline interference noise, it cannot be removed using notch filters. However, the cyclic property of the ECG signal can be used to extract only the desired wave. The moving average filter or the time-varying low pass filter \([104, 105]\) can be used to remove this type of noise.

4.2.2 QRS Detection

The detection of the QRS complex is one of the most fundamental steps in the fiducial-based ECG algorithms since it is considered a landmark of the heartbeat. However, the different states of cardiac diseases raised the need to QRS detectors that are able to correctly find the QRS under different morphologies. Not only that, the detector should also consider the noise removal algorithm used since some filters affect the complex, and consequently the detection accuracy.
Leif et al. [10] described the general workflow any the QRS detector as in figure 4.2. The role of each block can be described as follows:

- The linear filter: this usually utilize a band-pass filter to remove the unwanted frequencies components such as the P and T waves, while keeping the QRS complex.

- The non-linear transformation: the main role of this part is to enhance the QRS representation in the signal with respect to the background. However, not all the detectors implement this step.

- The decision rule: this part is to test whether the QRS complex exists or not. It could be implemented by a simple thresholding technique or more sophisticated algorithms.

**Pan and Tompkins Algorithm**

This algorithm was proposed to detect the QRS complex and was considered as an evolution in the ECG signal processing [106]. It follows the same workflow as in figure 4.2 utilizing the slope, amplitude, and width information of the signal points.

The signal is first filtered by a band-pass filter (sequence of low-pass and high-pass filters). This will highlight the QRS complex. Then, the filtered signal’s derivative is estimated and squared to highlight the QRS complex. Then, windows of QRS are determined from the ECG waveform. Finally, the R peaks are detected using adaptive thresholding.

We will employ this QRS detector to detect the R peaks in our approaches and consequently the heartbeats in the ECG signal. The reason behind this choice is that this filter is simple, consistent, accurate, and have high predictivity when compared to other detectors.

**4.3 ECG Features**

The challenge with the ECG signal is that we have 1D signals. Therefore, we can either extract the features from the time signal directly or add to those features the frequency-related information
with the expectation to increase distinction accuracy.

In this thesis, we utilize different ECG features. First, the 1D local binary patterns (LBP) feature extractor that extracts the features from the time-domain only. Second, in the time-frequency domain, two methods are used: the short-time Fourier transform as well as wavelet transform. This will allow us to gain more information about the frequency coefficients variation in the time signal. Finally, we utilize the nature of the ECG non-stationarity and use the cyclostationary function to describe the changes in the correlation of ECG signal over time.

4.3.1 1D Local Binary Pattern (LBP)

Local binary patterns (LBP) features are conventional, yet powerful, texture descriptors that have been used widely in 2D images \[107\]. However, Chatlani et al. \[108\] tweaked this 2D texture to propose a 1D LBP to be used on signals.

The main idea of the 1D LBP descriptor of the ECG signal is that we study 1xP neighbors of a certain sample point of the signal called the center point. This process is done by thresholding those neighbors based on the value of that center point \(c\). This thresholding results in a binary code of length equal to the number of neighbors studied. The equation for the LBP operator is illustrated below:

\[
LBP_p(c) = \sum_{r=0}^{p-1} \left\{ S \left[ I(c + r - \frac{P}{2}) - I(c) \right] 2^r + S \left[ I(c + r + 1) - I(c) \right] 2^{r + \frac{P}{2}} \right\}
\]  

(4.1)

where \(I\) is the signal intensity, and \(S\) is a sign function, that is, if the neighbor is greater than or equal to \(I(c)\), then it outputs one, and zero if otherwise. Figure 4.3 shows the operation of 1D-LBP on a synthetic signal. For each signal point, the binary code is converted into decimal and then stored in its place. Next, a histogram is built for each signal which accounts for the number of incidents of each code. Therefore, for \(P\) of 8, we have \(2^8\), or 256, possible codes including zero.

However, rotational invariant LBP \(r\) \[109\], an extension of the LBP, neglects the rotation information in the descriptor. Another extension of the LBP is the 'uniform LBP' (LBP\(_u\)), whereas uniform here is described as having the number of transitions less than 2 in the binary code. For instance, 01110111 has three transitions, thus, it is not uniform unlike 11101111. LBP\(_u\) stores all the non-uniform codes into one bin of the histogram resulting in \(2P(P-1)\) bins in the created histogram.

If both the uniform and rotational-invariant LBP are combined, a histogram of \(P+2\) bins can be created. In this thesis, we adapt the rotational-invariant uniform LBP\(_{u,r}\). The resulted histogram can be then used as features to describe the changes in the time-domain of the ECG signal.
4.3.2 Time-Frequency Features

In our time-frequency signal analysis, the Fourier transform as well as wavelet transform are utilized to gain more information about the frequency changes with time.

The utilization of the time-frequency properties is supported by the fact that ECG signal is a non-stationary signal. Therefore, the signal’s parameter can change over time. Thus, applying the Fourier transform to the whole signal or heartbeat will not reveal the frequency changes in those small intervals encompassed by the heartbeat.

Short-time Fourier Transform (STFT)

The main idea of the standard Fourier transform (FT) is to decompose the ECG signal into the frequencies that composes it. This can be done by performing an inner product of a family of basis functions with the ECG signal. These family bases are the complex oscillations $\exp(i\omega t)$, where $\omega$ is the frequency parameter. Therefore, the Fourier transform of a signal $x(t)$ can be given as follow:

$$\mathcal{F}\{x(t)\}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega \tau} x(\tau) d\tau \quad (4.2)$$

This transform can be interpreted as the frequency domain representation of the original signal. However, due to the fact that the ECG signal is a non-stationary signal, we can study the spectral behavior of the signal in short enough time window such that we can assume stationarity. This can be achieved by utilizing the short-time Fourier transform (STFT).

The STFT follows the same idea of the standard Fourier and adds the time dimension to the
base function by integrating a window of the complex exponential for the sake of localization to be \( \omega(t - t_0) \exp(i\omega t) \). Therefore, STFT of the ECG signal represents the Fourier coefficients of a fixed-size window contained within the signal. This can be achieved by sliding a window through the ECG time signal as shown in figure 4.4. An overlap between the windows is performed to avoid any discontinuities or artifacts in the windowing process. The mathematical formula of the STFT is as follow:

\[
STFT\{x(t)\}(\tau,\omega) = \int_{-\infty}^{\infty} \omega(t - \tau)e^{-i\omega \tau} x(\tau)d\tau
\]  

(4.3)

where \( \omega(t) \) is the window function.

There are different types of windows can be used such as rectangular, B-spline, Sine, Hann, or Hamming window. However, we utilize the Hamming window as it has better attenuation lobes and is given as:

\[
\omega(n) = a_0 - (1 - a_0)\cos\left(\frac{2\pi n}{N - 1}\right)
\]  

(4.4)

where \( N \) is the number of sample points in the window, and \( a \) is a constant. The shape of the window function as well as its Fourier is illustrated in figure 4.5.

When we apply this transformation on a 1D signal, it results in a time varying spectrum image, where one of the axes is the time and the other is the frequency. This image can be interpreted as
Figure 4.5: Hamming window function and its corresponding Fourier transform [5].

...the relative energy content changes in the frequency over time.

**Continuous Wavelet Transform (CWT)**

Although the STFT gives a good representation of the time-frequency characteristics of the signal, it suffers from the fixed resolution problem, which is not ideal in some situations. On the other hand, the wavelet transform solves this problem by providing a variable window size. This window size adapts to the frequency on the expense of the time resolution.

The main idea of the wavelets is to have a base function the same as the STFT. However, that base function has both localization and oscillation properties. Therefore, if the wavelet was scaled in the time, the oscillation frequency changes too.

The wavelet transform can be given as follow:

\[
W_\psi(t, s) = \int_{-\infty}^{\infty} \frac{1}{s^n} \psi \left( \frac{\tau - t}{s} \right) x(t) d\tau
\]

where \( x(t) \) is the time signal, \( s \) is the scale, and \( n \) is the scale normalization.

We utilize the generalized Morse wavelet [6, 110] in this thesis. This wavelet family can be considered as a superfamily for all commonly used analytic wavelets such as airy, Cauchy, complex exponentials, and Shannon. This helps to overcome the burden of spending time and effort of choosing with wavelet is better for which specific application. The generalized Morse wavelet family has two main parameters that controls the time and the frequency domains and is represented as follow:

\[
\Psi_{\beta, \gamma}(w) = \int_{-\infty}^{\infty} \psi_{\beta, \gamma}(t) e^{-i\omega t} dt = U(w) a_{\beta, \gamma}\omega^\beta e^{-\beta\gamma}
\]

where \( \omega \) is the frequency, \( \beta \) and \( \gamma \) are controlling parameters, \( U(w) \) is the unit step function, and
$a_{\beta, \gamma}$ is a normalization constant.

By changing the controlling parameters, $\beta$ and $\gamma$, the shape of the wavelet family is changed. The $\gamma$ controls the symmetry or skewness of the wavelet in time, or the high frequency decay in the frequency domain, while $\beta$ controls the behavior near zero in the time domain as in figure 4.6. Note that we have also the scale parameter that controls the compression or dilation of the wavelet.

The application of this transform on the 1D ECG signal yields a 2D image where one of the axes is the frequency and the other is the time.

### 4.3.3 Cyclostationary Signals

The majority of the current signal processing techniques for analyzing and describing signals usually use probabilistic models that are based on stationary statistics. This means that the probabilistic models do not vary with time. In such a way, the signal is represented or described using probabilistic parameters such as average, variance, or higher-order moments in a restricted small-time interval in order to hold the assumption of time-invariance of the parameters. However, this limits the information that can be extracted from the signal [111].

Nevertheless, cyclostationarity is a property of signal classes’ mathematical models where most of them are man-made modulated radio-frequency signals, such as cell phones, broadcast AM/FM/TV, satellites, Wi-Fi modems signals. The key of this property is that the probabilistic parameters of the signal vary with time either periodically or semi-periodically. Therefore, these parameters are defined for both time-domain signal and its frequency-domain representation, resulting in having
both temporal moments and spectral moments.

Suppose we have a continuous signal \( x(t) \), the autocovariance is the covariance at a pair of time points where the two series came from the same signal. The expected value or the mean \( E \) can be used to describe the autocovariance by:

\[
C_{xx}(t, t-1) = E[(x_t - \mu_t)(x_{t-1} - \mu_{t-1})] = Y_1
\]

\[
C_{xx}(t, t-2) = E[(x_t - \mu_t)(x_{t-2} - \mu_{t-1})] = Y_2
\]

\[
\vdots
\]

\[
C_{xx}(t, t-\tau) = E[(x_t - \mu_t)(x_{t-\tau} - \mu_{t-\tau})] = Y_\tau
\]

where \( \mu_t \) is the mean of the signal at time \( t \), \( \mu_{t-\tau} \) is the mean of the signal at time \( t-\tau \), and \( \tau \) is the lag. The value of \( Y_0 \) is called the variance where \( \tau = 0 \).

We use this parameter to quantify the relation of the signal \( x(t) \) with its past values. For instance, for a purely random process, the values of the signal have zero mean, and constant variances, which means that the signal’s autocovariance is zero.

**The Cyclic Autocorrelation Function (CAF)**

Equation 4.7 depends on the unit of \( x(t) \) (i.e. volt, millivolt, or even ampere). Therefore, we need to normalize it. We can perform this normalization by dividing by the variance to define the autocorrelation function (ACF).

\[
\hat{\rho}_\tau = \frac{Y_\tau}{Y_0}
\]

(4.8)

where \( \rho \) is the autocorrelation at lag \( \tau \) and takes values from -1 to 1 and determine the order of dependency, and \( Y \) is the autocovariance as mentioned before.

The mean value of the signal \( x(t) \) at a time point is given by:

\[
M_x(t, \tau) = E[x(t + \tau)]
\]

(4.9)

where \( t \) is the time point, and \( \tau \) is the lag.

This mean value is independent of the \( \tau \) in case of stationary signals [112]. Therefore, we have \( M_x(.) \) as:

\[
M_x(t, \tau_1) = M_x(t, \tau_2) = M_x(t, 0) = M_x(t)
\]

(4.10)

The ACF is the correlation between the random variables corresponding to two time instants of the
random signal (i.e. \(x(t_1)\) and \(x(t_2)\)). Thus, we have

\[
R_x(t_1, t_2) = E[ x(t_1)x^*(t_2) ]
\]  \hspace{1cm} (4.11)

where \(R_x\) is the ACF, * stands for the complex conjugate and \(t_1\) and \(t_2\) are two time instants of the signal \(x(t)\) as shown in figure 4.7.

The relation between the two time instants and the lag can be given as follows:

\[
t_1 = t + \frac{\tau}{2} \\
t_2 = t - \frac{\tau}{2} \\
t = \frac{(t_1 + t_2)}{2} \\
\tau = t_1 - t_2
\]  \hspace{1cm} (4.12)

Using these relationships, we can see how \(R_x\) changes with the time by defining the ACF in terms of the time and lag as follows:

\[
R_x(t, \tau) = E \left[ x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2}) \right] 
\]  \hspace{1cm} (4.13)

In case of stationary signals, this value of the autocorrelation depends only on the lag or separation between the two time points. Thus, the value of \(R_x\) is a function of \(\tau\) not \(t\).

\[
R_x(t_1, \tau) = R_x(t_2, \tau) = R_x(0, \tau) = R_x(\tau)
\]  \hspace{1cm} (4.14)
On the other hand, for non-stationary signals, the autocorrelation value depends on $t$. For instance, the cyclostationary signals can have its autocorrelation value either periodic or semi-periodic function. Having the fact that the autocorrelation varies with the time, we can represent it using a Fourier series [112].

$$R_x(t, \tau) = \sum_{\alpha} R_x^\alpha(\tau)e^{i2\pi\alpha t} \quad (4.15)$$

where $\alpha$ is the Fourier frequency and called the cycle frequency, or cyclic frequency (CF), and $R_x^\alpha(\tau)$ is the Fourier series coefficient and called the cyclic autocorrelation function. This cyclic autocorrelation function can be obtained in the usual manner of obtaining the Fourier coefficients

$$R_x^\alpha(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} R_x(t, \tau)e^{-i2\pi\alpha t} dt \quad (4.16)$$

using equation 4.13, we can obtain this cyclic autocorrelation function as follow:

$$R_x^\alpha(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t + \frac{\tau}{2}) x^*(t - \frac{\tau}{2})e^{-i2\pi\alpha t} dt \quad (4.17)$$

Note that we are talking about a second order cyclostationary signal. However, if we have a higher order signal for instance $n$, we would need $(n/2) + 1$ functions to properly characterize it.

The CAF can be seen as the amount of correlation between different frequency-shifted versions of the signal [111].

**The Spectral Correlation Function (SCF)**

The spectral correlation function (SCF) is most commonly used to characterize the cyclostationarity property. This is due to its computational efficiency since we need to compute the fast Fourier transform (FFT) that can be easily done on nowadays computers.

As the power spectrum function represents the spectral density of a time-averaged power [112], we can deal with the spectral correlation function the same way. SCF is considered as the spectral density or Fourier transform of the time-averaged correlation.

The motive behind this is that signals sometimes have unique features that are not clearly seen in the time domain. So, we need to investigate the frequency-domain representation. For instance, we can hardly detect a noisy sine way in the time-domain representation. However, if we converted it to the frequency-domain, we can easily detect the frequency components [111]. This is why we prefer to investigate the changes in the autocorrelation in the frequency domain.

In the same manner, the SCF of the cyclic autocorrelation function localize in the frequency
domain the amount of time-correlation between the frequency shifted versions of the signal \( x(t) \).

\[
S^\alpha_x = \int_{-\infty}^{\infty} R^\alpha_x(\tau) e^{-i2\pi\alpha t} d\tau
\]  

(4.18)

Therefore, the signal \( x(t) \) exhibits a second-order cyclostationary property, if the value of \( R^\alpha_x \) has non-zero values for some non-zero cyclic frequency \( \alpha \) [111].

Therefore, the value of the SCF will study the value of correlation of the signal at certain lag using different cyclic frequency as shown in figure 4.8. These changes can be used as distinctive features for the ECG signals.

4.4 Classifiers

In machine learning the classification process is to identify which category an observation belongs to. This can be done by a process called training. The training process teaches the classifier how to differentiate between the categories of the problem by giving it some instances that belongs to the classes of question. This can be utilized to either identify the type of heartbeat (i.e. either normal or arrhythmic), or to identify subjects using their ECG recordings in a biometric system.

There have been several classifiers proposed in the literature work of the machine learning such as K-nearest neighbors [113, 114], decision trees [115], and Bayes classifier [116, 117]. In this thesis, we will focus on support vector machine (SVM) and deep learning by utilizing the convolutional neural network (CNN).
4.4.1 Support Vector Machines (SVM)

A Support Vector Machine (SVM) is a supervised discriminative classifier, also called conditional classifier, that utilize the conditional probability $P(y|x)$ where $x$ is the input and $y$ is the observed label. The main idea of the SVM is to create a hyperplane that separates the classes. Given labeled training data, the algorithm outputs an optimal hyperplane that categorizes new examples as in figure 4.9. But what hyperplane is the best? This is what SVM choose based on the training data points.

![Figure 4.9: Support Vector Machine (SVM) data separation.](image)

Beside the idea of achieving the least number of misclassified instances from the classes, SVM tries to achieve a good margin. By good margin, we mean that the SVM tries to maximize the separation between the hyperplane and the classes’ points. This is the reason why SVM is sometimes called the maximal-margin classifier. The reason behind the maximization of the margin is to increase the confidence of classification. For instance, the red point in figure 4.9, if we chose line B, any changes can lead to misclassification of that point. On the other hand, if we chose line A, we can classify that point with higher confidence. The difference can be seen in the margin of the two hyperplanes, where hyperplane A has greater margin than B.

Let’s assume that we have linear binary classifier. By linear we mean that hyperplane is linear, while we have only two classes as output. Thus, $y \in \{-1, 1\}$. The classifier can be parameterized via this equation:

$$h_{\omega, b}(x) = g(\omega^T x + b)$$

(4.19)

where $h_{\omega, b}$ is the classification result, $x$ is the input, $\omega$ and $b$ are the parameters of the hyperplane, and $g(z)$ is a regression function such that if $z \geq 0$, then $g(z) = 1$, otherwise $g(z) = -1$.

If we have a training example $(x^{(i)}, y^{(i)})$, the function margin of the classifier with respect to the
training point is given by

$$\hat{\gamma}^{(i)} = y^{(i)}(\omega^T x + b) \quad (4.20)$$

Therefore, if we want to classify this example as $y^{(i)} = 1$, then the value of $\omega^T x + b$ should be large enough such that the regression result would be 1. Contrarily, if we want to classify this example as $y^{(i)} = -1$, then the value of $\omega^T x + b$ should be large enough negative number such that the regression result would be -1. Hence, the larger the marginal functional the greater confidence we have about the classification result.

Now, by considering the training set $S = \{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(n)}, y^{(n)})\}$ that has $n$ examples, we can describe the margin $\hat{\gamma}$ as the smallest functional margin among all the training examples in $S$ is shown in figure 4.10 and is given by

$$\hat{\gamma} = \min_{i=1,2,\ldots,n} \hat{\gamma}^{(i)} \quad (4.21)$$

![Figure 4.10: SVM margin for training set $S$](image)

However, to get the value of $\hat{\gamma}^{(i)}$, we notice that its direction is parallel to $\omega$. Thus, the value of point $B$ can be determined by $x^{(i)} - \hat{\gamma}^{(i)} \omega / ||\omega||$. Having this point $B$ on the hyperplane, it will satisfy the equation $\omega^T x + b = 0$. Therefore, we have

$$\omega^T \left( x^{(i)} - \hat{\gamma}^{(i)} \frac{\omega}{||\omega||} \right) + b = 0 \quad (4.22)$$
By solving for \( \hat{\gamma}^{(i)} \), we get

\[
\hat{\gamma}^{(i)} = y^{(i)} \left( \frac{\omega}{||\omega||} \right)^T x^{(i)} + \frac{b}{||\omega||} \]

(4.23)

The final margin for the training set can be then calculated using equation 4.21. As we discussed before, the SVM classifier tries to maximize this margin for the given training set. Therefore, the problem can be formulated as follow:

\[
\max_{\gamma, \omega, b} \gamma
\]

\[
such that \ y^{(i)} (\omega^T x + b) \geq \gamma, \forall i
\]

(4.24)

\[
||\omega|| = 1
\]

This is now an optimization problem that can be solved by different algorithms. Now, if we set the constraint on \( \hat{\gamma} \) and freed the constraint of \( ||\omega|| = 1 \), the problem would be easier and convex. Thus, finding a solution would be guaranteed. Thus, the new equation of the SVM classifier would be

\[
\max_{\gamma, \omega, b} \frac{1}{2||\omega||^2}
\]

\[
such that \ y^{(i)} (\omega^T x + b) \geq 1, \forall i
\]

(4.25)

By solving this equation, the values of \( \gamma \), \( \omega \), and \( b \) can be obtained during the training process. Having these parameters, a new example (i.e. testing set) can be assigned to either classes according to its features' values.

\subsection{Deep Convolutional Neural Network (CNN)}

In the normal neural networks, the main block is called neuron or cell. This network can be considered as a stack of layers, where each layer contains some neurons. In order to connect those layers, a set of weighted connections are created to connect the neurons of each layers as in figure 4.11. The convolutional neural network (CNN) is a deep neural network that has convolution layers which is inspired by the visual cortex.

The convolutional layer’s main role is to take the advantage of the 2D structure and pixel relations of the input image. For instance, the relation between pixel \((i,j)\) and its neighbor pixels in both dimensions. This can be achieved with local connections and weights of the kernels. Therefore, the weights of those convolution kernels are to be learned during the training process, where input images as well as their corresponding class labels are provided to the network. A backpropagation
algorithm [118] is performed to adjust those weights to achieve the lowest error provided in the loss function.

Those convolutional layers are usually followed by an activation layers in order to determine which neuron is activated. This can be done by using an activation function, usually a non-linear transformation function, that can be given as follow

\[ Y = Activation\left(\sum(weights \ast input) + bias\right) \]  \hspace{1cm} (4.26)

This activation function can be a binary step function, sigmoid, tanh, or rectified linear unit. However, a main concern when building such a function is the simplicity and convergence.

Another component of the CNN is the batch normalization layer. This layer allows the scaling of the activations in the convolution layer. In other words, it makes sure that there is no activation goes too high or too low.

The pooling layer is another component of the CNN. It allows to reduce the spatial size of the representation and consequently the parameters which will help in convergence.

Moreover, the dropout layer that deactivates some of the network’s neurons can be sometimes utilized in the network to avoid over-fitting during the training process.

However, the last layer is usually a softmax layer that has a number of outputs equal to the number of classes in the problem. It assigns a probability (i.e. value between 0 and 1) to each class, where the highest value will correspond to the assigned class.

The architecture of the CNN may differ from one application to another depending on the level of extraction and complexity of the features needed.
4.5 Performance Metrics

In order to verify a proposed approach, a performance measure should be provided that differs from one problem to another. For instance, in some classification problems the accuracy of the classifier can be enough indication of the performance. However, this measure is not enough in cases that the data is imbalanced. Moreover, some application such as the biometric-based systems, additional measures such as how many times trials the user will experience till he can access the system either being authorized or unauthorized.

Since we have two applications in this thesis (the medical application and biometric application) of the ECG signal, we will define a set of different measures that suites each approach.

In the arrhythmia detection, or detection of abnormal heartbeats from the ECG signal, we used the following performance measures:

1. Accuracy (Acc): It is the proportion of the correctly classified instances. We will use this measure in both the medical application (i.e. arrhythmia detection), and the biometric system. It can be defined as

\[ Acc_i = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \]  

(4.27)

where \( i \) is the class, and \( TP, TN, FP, \) and \( FN \) are the true positive, true negative, false positive, and false negative respectively.

2. Sensitivity (Sens): It measures the proportion of actual positives that are correctly identified as positive. We report this measure in the arrhythmia detection approach. It can be presented as:

\[ Sens_i = \frac{TP_i}{TP_i + FN_i} \]  

(4.28)

3. Specificity (Spec): The proportion of actual negatives that are correctly identified as negative, and is given as the following equation. We use this measure in the arrhythmia detection approach.

\[ Spec_i = \frac{TN_i}{TN_i + FP_i} \]  

(4.29)

4. F-score: It measures the system’s performance. However, it takes into consideration both the sensitivity and specificity. This measure is utilized in the arrhythmia detection approach.

\[ F - score_i = \frac{2 \cdot TP_i}{2TP_i + FP_i + FN_i} \]  

(4.30)
On the other hand, the biometric applications need other measures rather than only the accuracy of the system. These measures include:

1. Identification Rate (ID): It is the same as the accuracy. However, we use this term in the biometric approaches.

2. False Acceptance Rate (FAR): The probability that the biometric system will incorrectly accept an access attempt by an unauthorized user.

3. False Rejection Rate (FRR): The probability that the biometric system will incorrectly reject an access attempt by an authorized user.

4. Equal Error Rate (EER): It represents the error value at which false acceptances is equal to the proportion of false rejections.

5. The Cumulative Match Curve (CMC): the plot of the rank-k accuracy of the biometric system.

6. The Receiver Operating Characteristics Curve (ROC): the plot of the false acceptance rate (FAR) in the x-axis versus the genuine acceptance rate in the y-axis.

7. Area Under Curve the (AUC): this is the area under the ROC curve. It measures the discrimination ability of the system and it is equal to the probability that the system will rank a randomly chosen matching instance higher than a randomly chosen non-matching one.

**4.6 Summary**

In this chapter, we discussed the main methodology strategies that can be utilized when using the ECG signal to extract the information for further analysis. We started by discussing the types of databases and shed the light into the databases that we are using in this thesis. Then, we analyzed the noise source that may affect the ECG signal and how to remove it. Following that, we discussed the QRS complex detection from the ECG signal to extract the fiducial points. Afterwards, a detailed analysis was given of the ECG features such as the one-dimensional local binary pattern, short-time Fourier transform, continuous wavelets transform, as well as spectral correlation function, which we use as signal features. Thereafter, we gave a brief discussion on the support vector machine and neural network classifiers. Finally, the performance metrics for both clinical- and biometric-based systems were introduced.
Chapter 5

Automatic Detection of Arrhythmia

In this chapter, we will use the previously discussed methodology to implement a system that is able to detect abnormal heartbeats in the ECG signal of the patient.

5.1 Introduction

The challenge with the ECG signal is that it is 1D signal. Therefore, the features can be extracted using either the time domain or it can be further enriched using the time-frequency variations. We can achieve this by investigating the spectral features of the ECG signal which can be extracted using either the short-time Fourier transform, or the wavelet transform.

To the best of our knowledge, none of the approaches in the literature discussed before utilized the texture in ECG signal analysis.

In this work, we propose a methodological approach that combines the benefits of other proposed approaches to automatically classify different heartbeat conditions. We study the ability of the ECG signal to detect abnormalities using three different approaches. Having these approaches, we also study the utilization of different lead configurations. Lastly, we study the performance of these variations using two classifiers, so we can draw a conclusion on the best approach.

Although different types of cardiovascular abnormalities can be detected from the ECG signal, we focus on detecting normal and three different abnormal heartbeat types, including the left bundle branch block, the right bundle branch block, and paced beats.
Our approach is fully automated, and our primary contribution is in the novel structure of the feature extraction module, where we use textural information from both the temporal and spectro-temporal domain of the ECG signal.

5.2 Experimental Setup

The workflow of the proposed approach is illustrated in figure 5.1. After the acquisition of the ECG signal, it is preprocessed to remove the noise and prepare it to the following steps. The heartbeats are then extracted. Afterwards, the feature extraction takes place using three different approaches. Then, the classification of heartbeat is done using two different classifiers. Finally, the model validation is performed to study the generalizability.

![Workflow of the arrhythmia detection algorithm.](image)

**Figure 5.1:** Workflow of the arrhythmia detection algorithm.

### 5.2.1 Signal Preprocessing

In this work, we employ a bandpass of range 0.05-40 Hz. A notch filter is then used to remove the powerline noise at 60 Hz. Finally, we apply mean normalization on the signals such that we have a zero mean and standard deviation of 1.

### 5.2.2 Beats Extraction

After signal processing, we detect the QRS complexes and accordingly, the corresponding heartbeats. We employ the Pan and Tompkins (P&T) detector [71] to identify the R-peaks.

Afterwards, a constant number of points that account for 0.278 seconds is selected before and after each extracted R-point. Therefore, a final beat of 200 sample points or 0.556 seconds is resulted, which is considered sufficient to capture most of the heartbeat information [39].

### 5.2.3 Feature Extraction

In this work, we employ the 1D local binary pattern (LBP) to extract the features in the time domain, as well as the time-frequency domain features using short-time Fourier transform (STFT)
and wavelets or precisely the generalized Morse wavelet transform (CWT).

**One-dimensional Local Binary Pattern**

After segmenting the ECG heartbeats, we employed the previously discussed 1D LBP to extract the one-dimensional textures in the time domain. In this work, we utilized the uniform rotational invariant LBP.

Moreover, we studied the effect of the size of studied neighbors $P$ on the accuracy of classification of the abnormality of the ECG.

![Figure 5.2: The STFT images of the four beats types. The second row is the difference between each heartbeat type image and the normal image. The third row is the GLCM segments images for each beat [4].](image)

**Short-time Fourier Transform (STFT)**

Since STFT describes the information of the frequency variation in time of the ECG signal, we can use it to distinguish the ECG signal characteristics.

In order to study these spectral variations, the STFT image is sliced into three segments along the frequency domain. Afterwards, for each segment we use two texture operators: the 2D local binary patterns, and the Haralick features [119] of the grayscale co-occurrence matrix (GLCM) such as the entropy, homogeneity, and contrast.

Figure 5.2 shows the difference between the three types of heartbeats. To emphasize the difference, we subtracted the STFT image of each heartbeat from the normal one. In addition, we illustrated the GLCM images of each segment of the three segments.
Continuous Wavelet Transform (CWT)

Following the same manner, the CWT image is divided into three segments and the 2D LBP and Haralick features are extracted from each segment. Figure 5.3 shows the difference in the images' properties of the four heartbeat types.

![CWT Images](image)

Figure 5.3: The CWT images of the four beats types. The second row is the difference between each heartbeat type image and the normal image. The third row is the GLCM segments images for each beat [4].

5.2.4 Classification and model validation

In the classification process, we use Support Vector Machines (SVM) [120] as well as the K-nearest neighbor (KNN) classifiers for comparison. We utilize cross-validation to potentially eliminate any over-fitting problems. The performance measures in this work include the accuracy, sensitivity, specificity, and the F-score.

5.2.5 Lead Configurations

In this work, we tested the classification of heartbeats from different ECG leads. We employed the limb lead (Lead II) alone, lead V alone, and combining the two leads to investigate the effect of each lead on the abnormality detection.
5.3 Results

We employed the texture of the ECG signal in both temporal and spectro-temporal domains. Each signal is filtered as discussed before. The R peaks detection was carried out as shown in figure 5.4 with an accuracy of 99% when compared to the labels that accompanied the database.

![Figure 5.4: R peak detection in the ECG signal of one of the records in MITDB.](image)

We only utilize the MITDB database for testing in this work. However, since the four classes that we use are of different distribution (i.e. different number of instances), we downsample the classes that encompasses large number of instances by changing the ratio involved in the training process. Table 5.1 shows the training distribution of each class. However, in order to avoid any bias in the training process, 10-fold cross validation is performed.

<table>
<thead>
<tr>
<th>Beat Type</th>
<th>Total #</th>
<th>Training #</th>
<th>Testing #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>74762</td>
<td>7476</td>
<td>67286</td>
</tr>
<tr>
<td>Left bundle block</td>
<td>8075</td>
<td>5652</td>
<td>2423</td>
</tr>
<tr>
<td>Right bundle block</td>
<td>7259</td>
<td>5081</td>
<td>2178</td>
</tr>
<tr>
<td>Paced</td>
<td>7028</td>
<td>4919</td>
<td>2109</td>
</tr>
</tbody>
</table>

A kernel size of 5.7 was used in the radial basis function of SVM, while the number of neighbors in the KNN classifiers are 5 utilizing the Euclidean distance.

The 1D-LBP was implemented, and the effect of different neighbor sizes $P$ was studied. This performance variation can be shown in figure 5.5 having the accuracy values starting at 0.6.

As can be seen in the figure, the accuracy increases with increasing the P value. This is expected since we gain more information about the signal while increasing the number of neighbors involved in the calculations around the center points.

As we mentioned before that we compared different classifiers with different lead configurations,
Figure 5.5: Effect of P size on the testing data for both classifiers using the two ECG leads, where the SVM performance is in solid, and KNN is in dotted lines [4].

tables 5.2, 5.3, and 5.4 shows the performances of each experiment.

Table 5.2: LBP texture feature performance on the MITDB dataset.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Performance</th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lead V</td>
<td>Lead II</td>
</tr>
<tr>
<td>SVM</td>
<td>Acc.</td>
<td>84.89</td>
<td>88.96</td>
</tr>
<tr>
<td></td>
<td>Sens.</td>
<td>86.5</td>
<td>88.04</td>
</tr>
<tr>
<td></td>
<td>Spec.</td>
<td>96.9</td>
<td>96.44</td>
</tr>
<tr>
<td></td>
<td>F-score</td>
<td>0.8789</td>
<td>0.8748</td>
</tr>
<tr>
<td>KNN</td>
<td>Acc.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Sens.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Spec.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>F-score</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Using the LBP features, we can notice that the KNN classifier over-fitted during the training. For instance, it reached 100% accuracy in training and 78.48% in testing using the two leads. Whereas the SVM resulted in a better performance when compared to KNN by taking into consideration both training and testing using the two ECG leads.

In the STFT texture features, we used a window of size 30 with an overlap size of 20 to examine the frequency range of 1:180 Hz. A 2D-LBP extractor of $P=8$ and the GLCM features were used for each image segment.

The SVM classifier achieved a higher performance of 99.24% when compared to the KNN that achieved 96% accuracy when utilizing both ECG leads. Moreover, when only one ECG lead was utilized, we achieved an SVM accuracy of about 96% in the SVM compared to 90.16% in KNN. Therefore, we can conclude that SVM classifier exhibit higher performance independent of the texture used.

We can also notice that the two single lead configurations (i.e. lead II only and lead V only)
Table 5.3: STFT texture feature performance on the MITDB dataset.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Performance</th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lead V</td>
<td>Lead II</td>
</tr>
<tr>
<td>SVM</td>
<td>Acc.</td>
<td>96.93</td>
<td>96.57</td>
</tr>
<tr>
<td></td>
<td>Sens.</td>
<td>98.33</td>
<td>99.21</td>
</tr>
<tr>
<td></td>
<td>Spec.</td>
<td>99.79</td>
<td>99.92</td>
</tr>
<tr>
<td></td>
<td>F-score</td>
<td>0.9963</td>
<td>0.9960</td>
</tr>
<tr>
<td>KNN</td>
<td>Acc.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Sens.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Spec.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>F-score</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

showed comparable performance results using both SVM and KNN. This is due to the fact that normal beats are usually distinguishable in the first lead. Whereas the ectopic beats will be discriminated more efficiently in the second lead. However, it is difficult to distinguish normal beats in the second lead due to the fact that this axis is nearly orthogonal to the mean cardiac electrical axis.

In the CWT on the other hand, we used values of 60 and 3 for $\beta$ and $\gamma$ respectively. Therefore, a wavelet shape that lies between the Airy wavelet and derivative of Gaussian were utilized. A scale parameter of 0 to 47 was also used in the wavelet function. In the 2D LBP, we used the same value of $P=8$. The resulted feature vector from the 2D LBP is then joined to the GLCM features to create one feature vector. Again, the SVM achieved higher accuracy than KNN, with a training and testing accuracies of 99.87% and 99.81% respectively.

Table 5.4: CWT texture feature performance on the MITDB dataset.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Performance</th>
<th>Training</th>
<th>Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lead V</td>
<td>Lead II</td>
</tr>
<tr>
<td>SVM</td>
<td>Acc.</td>
<td>99.01</td>
<td>98.99</td>
</tr>
<tr>
<td></td>
<td>Sens.</td>
<td>98.70</td>
<td>98.80</td>
</tr>
<tr>
<td></td>
<td>Spec.</td>
<td>99.91</td>
<td>99.90</td>
</tr>
<tr>
<td></td>
<td>F-score</td>
<td>0.9907</td>
<td>0.9915</td>
</tr>
<tr>
<td>KNN</td>
<td>Acc.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Sens.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Spec.</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>F-score</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

As mentioned before that we performed 10-fold cross-validation to avoid any bias or aver-fitting. Therefore, we need to analyze the performance each validation iteration. Figure 5.6 shows the accuracies of the ten cross-validations. Therefore, in this work, the overall performance represents the average for across all the validation iterations.

As can be seen in the cross-validation figure, the variance between the performances in each configuration is not high except for LBP using only one lead. However, when utilizing both leads,
we always achieve higher accuracy than only one lead. This can be seen in the three approaches (LBP, STFT, and CWT). On the other hand, while STFT outperforms the LBP, CWT has the highest performance for all leads configurations.

So far, we investigated the performance of the overall system. Nevertheless, we need more insight on the performance of each class (i.e. heartbeat type) to further investigate the behavior of each system.

Figure 5.7 shows the performance of each heartbeat type using the three texture features with different leads configurations and SVM classifier. Although the 1D-LBP shows high accuracy, sensitivity, and specificity measures, the F-score is low for the abnormal classes. However, the F-scores of the STFT is higher than the 1D-LBP. This is because we integrate the frequency information into our features domain. On the other hand, the CWT exhibits high values of all measures including
the F-scores of all ECG beats as it utilizes different window resolution that outperform the STFT.

To compare the propose approach with the literature work, table 5.5 shows the properties and performances of our approach and the intra-patient approaches in the literature using the MITDB database.

Table 5.5: Comparison of the proposed approach with intra-patient approaches in the literature using the MITDB database.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Seg.</th>
<th>Features</th>
<th>Classifiers</th>
<th>Acc.</th>
<th>Cls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Butt et al. [43]</td>
<td>Auto</td>
<td>Statistical features from temporal domain</td>
<td>SVM</td>
<td>98.78</td>
<td>15</td>
</tr>
<tr>
<td>Dewangan et al. [44]</td>
<td>Auto</td>
<td>Morphological features in temporal domain &amp; wavelet coefficients</td>
<td>ANN</td>
<td>78.01</td>
<td>6</td>
</tr>
<tr>
<td>Assadi et al. [39]</td>
<td>—</td>
<td>Frequency features</td>
<td>KNN</td>
<td>95.31</td>
<td>3</td>
</tr>
<tr>
<td>Tran et al. [121]</td>
<td>—</td>
<td>Temporal Hermite basis functions and RR-interval</td>
<td>Ensemble</td>
<td>98.63</td>
<td>7</td>
</tr>
<tr>
<td>Evaluated LBP</td>
<td>Auto</td>
<td>Texture in temporal domain</td>
<td>SVM</td>
<td>92.97</td>
<td>4</td>
</tr>
<tr>
<td>Evaluated STFT</td>
<td>Auto</td>
<td>Texture in spectro-temporal domain</td>
<td>SVM</td>
<td>99.24</td>
<td>4</td>
</tr>
<tr>
<td>Proposed CWT</td>
<td>Auto</td>
<td>Texture in spectro-temporal domain</td>
<td>SVM</td>
<td>99.81</td>
<td>4</td>
</tr>
</tbody>
</table>

Seg.=Segmentation, Cls=Classes Number.

As show, the proposed system outperformed other approaches discussed in the literature. Nevertheless, the comparison is not completely fair since we use all the patients in the database including the noisy records and patients that have a pacemaker on and the fact that we use cross-validation with different training-testing ratio and report the average accuracies. For instance, Dewangan et al. [44] used some selected patients instead of using the whole patients in database, which can enhance the results. On the other hand, although M. Butt et al [43] used the whole 15 classes, the unbalance in data distribution and the difference in the number of instances per class was so great that can make the grid search for the classifier parameters over-fit (i.e. the normal class has more than 70,000 instances, while some of the classes have less than 100 instances).

5.4 Summary

In this chapter, we proposed our approach to automatically detect arrhythmia in ECG signals by utilizing signal textures. We employed the 1D local binary pattern (LBP) to study the signal texture in the time domain. We also used the spectro-temporal domain textures by utilizing the short-time Fourier transform (STFT) and generalized Morse wavelet transform (CWT). Different classifiers and lead configurations were investigated. The results showed that utilizing two ECG leads enhanced the accuracy. Moreover, the SVM classifier outperformed the KNN that over-fitted. The proposed approach using the CWT features and SVM classifier achieved an accuracy of 99.8%.
Chapter 6

ECG-based Human Authentication using High-level Spectro-temporal Features

In this chapter, we will propose an approach that uses the spectro-temporal images of the ECG signals that discussed before to identify and verify subjects in a biometric system.

6.1 Introduction

The spectro-temporal features of the ECG signal can also be used in distinguishing individuals' characteristics. Therefore, we can build an ECG-based biometric system by utilizing these feature spaces. We believe that literature work in the ECG biometric did not address the utilization of deep learning along with the spectro-temporal features of the ECG signal, specifically convolutional neural network. In this approach, a high-level representation of spectro-temporal features is extracted using the intermediate layers of the convolutional neural network (CNN).

However, Camara et al. [74] and Bassiouni et al. [73] combined the frequency and time domain features, but they only used separate features from each domain. Moreover, Zhang et al. [86] used both deep learning and wavelets-based approach, but different combination of scales had to be selected before the classification takes place base on the value of the final identification rate. Pouryayevali et al. [122] utilized the wavelet of the ECG signal, however, a feature selection was performed prior to the classification using the linear discriminant analysis.
6.2 Experimental Setup

The workflow of our approach is illustrated in figure 6.1. We first start with the data acquisition, then we process the acquired data. A feature extraction then is performed by utilizing the STFT and CWT images. A CNN is used to extract the high-level features followed by both identification and verification procedures. Finally, model robustness and generalization are tested using both cross-validation and different databases.

![Workflow Diagram](image)

Figure 6.1: Workflow of the ECG biometric approach using high-level spectro-temporal features.

6.2.1 Signal Preprocessing

In this work, we first resample all the data to 360 Hz since we use different databases. We then apply a bandpass filter with a range of 0.5-40 Hz, and a notch filter of 60 Hz to remove the powerline interference. The R peaks are detected using the Pan and Tompkins [71], then segments of 0.56 seconds are extracted using the R peaks (99 sample point before, and 100 after the R peak). We utilized the limb lead (i.e. lead II) since it is common it all the databases.

6.2.2 Feature Extraction

Following the same algorithm of extracting the STFT and CWT images from the signal, we create the spectral images. Figure 6.2 shows the intra- and inter-subject variability using the spectral images of two subjects from the MITDB database.

Two subjects were randomly picked from the MITDB database, then the spectro-temporal images using STFT and CWT are created for an average of randomly selected 10 heartbeats. These images were created two times for each subject, each is for different averaged 10 heartbeats. Thus, a total of eight images are created, four for each subject, each is for different 10 heartbeats. We subtract the same type of images of the same subject (i.e. STFT or CWT) to study the intra-subject variability, while the inter-subject variability can be studied by subtracting the images of the same type (i.e. STFT or CWT) of the two different subjects.
Figure 6.2. The intra- and inter-subject variability of both STFT (left) and CWT (right) images of two subjects from the MITDB [9].

As can be seen from the figure, this small intra-subject variability and high inter-subject variability resulted in the STFT and CWT motivated us to utilize these spectral-temporal images to feed the convolutional neural network in the proposed ECG biometric system.

6.2.3 Individual Identification and Verification

Deep learning has gained the attention recently because it allows the extraction of high-level features that are able to discriminate the individual with high accuracy. In this approach, we utilize a 2D deep convolutional neural network to extract those features from the spectro-temporal domain images.

The architecture of the proposed network is shown in figure 6.3, whereas the size of each layer is in table 6.1. The input of the network is the spectro-temporal images (STFT or CWT) of the heartbeat. As we can notice, the size of the ID-vectors in this architecture is 100. Therefore, a representation compression is achieved since we represent 200 data point (i.e. the heartbeat) with 100 highly discriminative features. Thus, a database with less memory can be created for the gallery set.
Figure 6.3. The architecture of the proposed CNN using the spectral image as input and the extracted feature to be compared with other subjects for identification.

Table 6.1. The proposed CNN layers information.

<table>
<thead>
<tr>
<th>Layer ID</th>
<th>Layer Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input Layer</td>
<td>128x128</td>
</tr>
<tr>
<td>Conv_1</td>
<td>5x5@16</td>
</tr>
<tr>
<td>Conv_2</td>
<td>5x5@32</td>
</tr>
<tr>
<td>MaxPool</td>
<td>2x2</td>
</tr>
<tr>
<td>FC</td>
<td>100</td>
</tr>
<tr>
<td>Softmax</td>
<td>Variable</td>
</tr>
</tbody>
</table>

6.2.4 Model Validation

In this approach, we validate the model on eight public ECG databases. In each database, we divide the signal into training and testing sets. In the training phase, both the images and the labels are used to train the model (i.e. supervised learning). However, in the testing phase, the testing set is subdivided into gallery and probe sets.

The gallery set represents the reference vectors of the individuals that are stored in the database, where each vector is called template (each subject has his own template), while the probe is a simulation of the data acquired from the subject during the authentication. This probe data is tested against the gallery. In this work, the gallery and probe vectors represent the high-level features that are extracted from the fully connected layer.

In order to measure the similarity between the probe and template vectors, a distance metric is adopted such as the Euclidean distance. However, due to the variability that can be found in the ECG signal, we average M heartbeats feature vector that belongs to the same subject and use it as template vector.

Moreover, we adopt cross-validation to avoid any bias or over-fitting. Therefore, the training and testing of the system is done on the base of 5-fold cross-validation. In addition, the selection of the
gallery and probe in the testing set is repeated 10 times each with a random selection. This results in a total of 5x10=50 result records. Therefore, the results that we will discuss in this chapter is the average of those 50 iterations.

In this approach we use different metrics to assess the model including the identification rate (accuracy), the CMC curve, the ROC curve, and the AUC measures.

6.3 Results

In this work, we tested the proposed approach on eight public ECG databases. These databases have both normal and abnormal heartbeats (i.e. ventricular flatter, left or right bundle block, atrial fibrillation, etc.) as illustrated in table 4.1. We utilized only 15 minutes of each recording for both training and testing combined.

Each database is then divided into training and testing of 80% and 20% respectively. A learning rate of 0.002 was used in the training process. In each validation iteration, we averaged 10 randomly selected vectors to act as the gallery set, while the probe set represent another 100 randomly records.

In the identification scenario, the CMC curves of both approaches are illustrated in figure 6.4. The VFDB scored the least average identification rate in both algorithms with a value of approximately 90%. However, it raised by 7% in rank-2 exceeding 97%. On the other hands, all other databases identification rates are above 97% at rank-1.

Although the performance of the both STFT and CWT features are close, the CWT achieved higher identification rate in the CEBSDB database with a difference of 0.8%. Therefore, both approaches provide a good performance in the identification procedure, while having the STFT outperformed the CWT overall the databases.

In the verification scenario, we concerned about the EER, ROC curve, and AUC. The EER of STFT is greater than CWT in PTBDB, while it is equal in performance in VFDB and has a lower EER in the rest of the databases as shown in figure 6.5 (a).

Although the average EER of both the STFT and CWT are 0.0267, and 0.03859 respectively, the maximum EER achieved by both approaches is 0.0629 at the VFDB. Therefore, we can consider that the STFT is slightly better than the CWT in terms of EER.

The ROC curves of both approaches can also be used to assess and compare both algorithms. Figure 6.6 shows the ROC of both STFT and CWT features. In these curves, we can notice that all databases have high performance, except for the VFDB since as we mentioned before has a sever ventricular diseases that affects the QRS complex.
Figure 6.4. The CMC curve of the proposed approach of both STFT and CWT using ECG different databases [9].

However, to investigate more deeply in the ROC curves, we estimated the AUC for both algorithms. As shown in figure 6.5, The STFT has a better performance than the CWT in the normal databases, while this performance varies in the abnormal databases.

The average AUC of both STFT and CWT are 0.993, 0.988 respectively. Therefore, we can notice that the STFT has a better performance as features on average when compared to CWT. This can be explained based on our previous arrhythmia approach in the previous chapter, in which we proved that CWT sustain the discriminative ability of the system to detect arrhythmia heartbeats when compared to STFT. This in turn leads to higher intra-subject variability, and therefore lower discriminative power in the biometric system (i.e. inter-subject variability).

To summarize the performance of each approach, table 6.2 shows the identification rate (ID), EER, AUC for each approach and databases. From the table, the normal databases achieved an average identification rate and EER of 99.23% and 0.041 respectively using the STFT. This result
is close to the CWT results, however, the CWT has lower identification rate and AU, while keeping higher EER.

Contrarily, the abnormal databases had equal values of AUC for both approaches. Moreover, the difference the STFT and CWT average identification rate is higher around 0.5%, whereas it is 0.2% in normal databases. Overall the databases, the EER for the STFT is lower compared to CWT. The STFT approach achieved a total of 97.85% overall the eight databases, which is also close to the CWT. However, it achieved lower average EER, and higher AUC.

Table 6.3. Comparison of the average performance of proposed approach with the literature work.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Databases (count) Type</th>
<th>Average IDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yue et al. [123]</td>
<td>(3) N/AbN</td>
<td>85.1</td>
</tr>
<tr>
<td>Dar et al. [87]</td>
<td>(3) N</td>
<td>93.2</td>
</tr>
<tr>
<td>Agrafioti et al. [37]</td>
<td>(3) N</td>
<td>96.2</td>
</tr>
<tr>
<td>Tantawi et al. [124]</td>
<td>(3) N</td>
<td>95.9</td>
</tr>
<tr>
<td>Zhang et al. [86]</td>
<td>(8) N/AbN</td>
<td>93.5</td>
</tr>
<tr>
<td>Evaluated Approach (CWT)</td>
<td>(8) N/AbN</td>
<td>97.5</td>
</tr>
<tr>
<td>Proposed Approach (STFT)</td>
<td>(8) N/AbN</td>
<td>97.85</td>
</tr>
</tbody>
</table>

, N=Normal, AbN=Abnormal, IDR= Identification Rate.
We also compared the performance of this approach with other work in the literature as shown in table 6.3. As shown, both of the proposed approaches achieved the highest performance when compared to the literature work. Although validating using less dataset may increase the performance, the proposed approach using spectro-temporal features achieved 97.9% average accuracy on different eight databases using the STFT images as input to the convolutional neural network. The utilization of abnormal heartbeats is an advantage since it impersonate the real-world scenario as the subject might suffer from any cardiac condition.

6.4 Summary

In this chapter, we proposed an ECG-based biometric system that uses the spectro-temporal images to feed a convolutional neural network. The features were extracted from the fully connected layer of the network, then a Euclidean distance was used to measure the similarity. Eight databases as well as cross-validation were used to validate the proposed model. An average identification rate of 97.85% and 97.5% were achieved using both STFT and CWT images respectively.
Chapter 7

ECG-based Human Identification using Spectral Correlation

In this chapter, we will propose an approach that uses the cyclostationary property of the ECG signals that was explained before to study the distinctive characteristics of individuals to identify and subjects in a biometric-based system.

7.1 Introduction

In this work, we propose a different ECG biometric-based approach to identify subjects. We utilize the cyclostationary property of the ECG signal, namely the spectral correlation function (SCF), in which the properties of the signal change with time either periodically or semi-periodically.

As mentioned before, the conventional workflow of the biometric system is 1) data acquisition, 2) data preprocessing, 3) feature extraction, 4) user identification, 5) model validation.

However, this approach has several advantages over the previously proposed biometric system.

1. It does not require the detection of the R peaks of the ECG beats. Thus, it is more robust to noise effect. This method is called blind segmentation method since no fiducial points required.

2. It does not utilize the noise removal. This is due to the fact that noise exhibit a non-cyclostationary properties. Thus, it will not affect the algorithm since we use the cyclostationarity of the signal.
7.2 Experimental Setup

The workflow of this approach is illustrated in figure 7.1. The workflow is similar to the previously proposed biometric approach except for the filtration and extraction of R peaks.

**Figure 7.1.** Workflow of the ECG biometric approach using the cyclostationary properties.

7.2.1 Signal Preprocessing

In this approach, we first resampled the data to 360 Hz since we use different databases too. A blind segmentation of two-second segments of the ECG signal was performed. These segments are not overlapped to correctly validate the algorithm.

7.2.2 Feature Extraction

In this approach we used the aforementioned spectral correlation properties of the ECG signal. The motivation behind that is that these cyclostationarity can differ from one subject to another.

Figure 7.2 shows the intra- and inter-subject variability between the SCF image of two subjects of the MITDB. As can be seen, the difference between the images that belong to the same subject is low, which implies low intra-subject variability. On the other hand, the difference between the SCF of two different subjects is high, therefore, the inter-subject variability is high. Thus, we can use these functions, the spectral correlation function, as a biometric feature that can distinguish individuals.

7.2.3 Individual Identification

We utilize the same architecture of the CNN as in the previous biometric approach. However, in this approach, we use the output of the softmax layer as it was a classification problem. The size of each layer is shown in table 6.1. Again, the softmax layer’s size is variable according to the number of subjects in each ECG database.
Intra-subject Variability

Inter-subject variability

Figure 7.2. The intra- and inter-subject variability in the spectral correlation images of averaged 10 ECG segments. Each row represents two different groups of 10-averaged ECG segment per one subject and their difference to show the intra-subject variability. Whereas columns represent the difference between the images between different subjects in MITDB database.

### 7.2.4 Model Validation

In order to test the robustness and generalizability of the proposed approach, we utilize different biometric measures such as the identification rate, CMC curve, FAR, and FRR as illustrated before. We will test this approach on seven different databases, namely: MITDB, CEBSDB, NSRDB, STDB, Fantasia, AFDB, and VFDB as in table 4.1 that encompasses both normal and abnormal heartbeats.

### 7.3 Experimental Results

We utilized different duration for each database in our approach. For instance, we used all the acquired signals from all databases of interest except for Fantasia and AFDB where we used only 60 minutes.
We utilized the limb lead (i.e. lead II) since it is common in all the databases. Then a non-overlapping blind segmentation of two-second segments was done. Thus, a total of 720 sample points is each ECG segment are used for identification. Afterwards, we created the spectral correlation images for each 2-second ECG segment.

The resulted images are then resized to 128x128 to feed the CNN proposed for the identification process. A learning rate of 0.002, and 15 epochs were used to train the network on the training set using stochastic gradient descent algorithm and backpropagation algorithm.

In the model validation, we divided each database into 80% training and 20% testing, and repeated this for 5-fold cross-validations to measure the robustness of the proposed approach. Figure 7.3 shows the cross-validation results for each database. As shown in the figure, all the dataset had a median identification rates higher than 90%, expect for the VFDB that achieved a median of 86.3%, while having the CESDB achieved an identification rate of 100%.

![Figure 7.3. The boxplot of 10 validation experiments on each dataset using the proposed approach.](image)

The CMC curve of the proposed system is illustrated in figure 7.4. The CEBSDB achieved the highest identification rate of rank-1 of 99.6%, and it reaches 100% in rank-2. In the MITDB, the identification rate of rank-1 is 96.5%, while it reaches 99.5% at rank-5. Whereas NSRDB starts with an average identification rate of 98.7% and reaches 100% in rank-3. On the other hand, STDB, Fantasia, and AFDB databases reach 100% rate at rank-10. Whereas the VFDB has a
steeper ascending curve before rank-2. However, all the databases tested achieved more than 98% identification rate at rank-5.

![Figure 7.4. The CMC curve of the proposed approach on the validation databases.](image)

We summarized performance of the databases in table 7.1 compared to the literature work. Beside the identification rate, we reported the average FAR and FRR rates to gain more information about the performance. As shown in the table, the FAR is 0.4% for both normal and abnormal databases. On the other hand, the FRR is zero in normal databases, while having a value of 0.03% in abnormal databases. The overall FAR and FRR across all the databases are 0.4% and 0.02% respectively.

Table 7.1: The proposed approach’s average identification rates on different ECG databases.

<table>
<thead>
<tr>
<th>Database</th>
<th>Identification Rate</th>
<th>FAR</th>
<th>FRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEBSDB</td>
<td>0.996</td>
<td>0.008</td>
<td>0.00</td>
</tr>
<tr>
<td>NSRDB</td>
<td>0.987</td>
<td>0.003</td>
<td>0.00</td>
</tr>
<tr>
<td>Fantasia</td>
<td>0.945</td>
<td>0.114</td>
<td>0.00</td>
</tr>
<tr>
<td>Average Normal</td>
<td>0.974</td>
<td>0.040</td>
<td>0.00</td>
</tr>
<tr>
<td>MITDB</td>
<td>0.965</td>
<td>0.001</td>
<td>0.00</td>
</tr>
<tr>
<td>STDB</td>
<td>0.972</td>
<td>0.019</td>
<td>0.00</td>
</tr>
<tr>
<td>AFDB</td>
<td>0.944</td>
<td>0.100</td>
<td>0.01</td>
</tr>
<tr>
<td>VFDB</td>
<td>0.826</td>
<td>0.04</td>
<td>0.00</td>
</tr>
<tr>
<td>Average Abnormal</td>
<td>0.927</td>
<td>0.040</td>
<td>0.003</td>
</tr>
<tr>
<td>Average All</td>
<td>0.948</td>
<td>0.040</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Finally, the proposed approach is compared to other approaches in the literature as shown in table 7.2. Our approach shows its superiority in three ways, either to use less ECG segment duration as input, studying subject with abnormal ECG signals, or the blind segmentation of the signal. For
instance, some approaches use the healthy subjects in the ECG database to test their algorithm. This increases in the identification rate since ECG abnormality may affect the intra-subject variability.

Table 7.2: Comparison of the proposed approach with the literature work on the databases level.

<table>
<thead>
<tr>
<th>Database</th>
<th>Authors</th>
<th>Subjects</th>
<th>Blind Segmentation</th>
<th>Required ECG time</th>
<th>Acc. %</th>
</tr>
</thead>
<tbody>
<tr>
<td>MITDB</td>
<td>Bassiouni et al. [73]</td>
<td>30</td>
<td>Yes</td>
<td>10 seconds</td>
<td>96.6</td>
</tr>
<tr>
<td></td>
<td>Shen et al. [72]</td>
<td>20</td>
<td>No</td>
<td>20 beats</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Dar et al. [87]</td>
<td>47</td>
<td>No</td>
<td>——</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Tang et al. [125]</td>
<td>10</td>
<td>No</td>
<td>——</td>
<td>91.7</td>
</tr>
<tr>
<td></td>
<td>Zhang et al. [86]</td>
<td>47</td>
<td>Yes</td>
<td>2 seconds</td>
<td>91.1</td>
</tr>
<tr>
<td></td>
<td>Proposed Approach</td>
<td>47</td>
<td>Yes</td>
<td>2 seconds</td>
<td>96.5</td>
</tr>
<tr>
<td>CESDB</td>
<td>Zhang et al. [86]</td>
<td>20</td>
<td>Yes</td>
<td>2 seconds</td>
<td>99.0</td>
</tr>
<tr>
<td></td>
<td>Proposed Approach</td>
<td>20</td>
<td>Yes</td>
<td>2 seconds</td>
<td>99.6</td>
</tr>
<tr>
<td>NSRDB</td>
<td>Venkatesh et al. [70]</td>
<td>15</td>
<td>No</td>
<td>1 minute</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Camara et al. [74]</td>
<td>18</td>
<td>Yes</td>
<td>2 seconds</td>
<td>94-96.67</td>
</tr>
<tr>
<td></td>
<td>Dar et al. [87]</td>
<td>18</td>
<td>No</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zhang et al. [86]</td>
<td>18</td>
<td>Yes</td>
<td>2 seconds</td>
<td>95.1</td>
</tr>
<tr>
<td></td>
<td>Fatemian et al. [77]</td>
<td>13</td>
<td>No</td>
<td>2 beats</td>
<td>99.6</td>
</tr>
<tr>
<td></td>
<td>Palaniappan et al. [67]</td>
<td>10</td>
<td>No</td>
<td>——</td>
<td>97.6</td>
</tr>
<tr>
<td></td>
<td>Agrafioti et al. [37]</td>
<td>13</td>
<td>Yes</td>
<td>10 seconds</td>
<td>96.2</td>
</tr>
<tr>
<td></td>
<td>Proposed Approach</td>
<td>18</td>
<td>Yes</td>
<td>2 seconds</td>
<td>98.7</td>
</tr>
<tr>
<td>STDB</td>
<td>Zhang et al. [86]</td>
<td>28</td>
<td>Yes</td>
<td>2 seconds</td>
<td>90.3</td>
</tr>
<tr>
<td></td>
<td>Proposed Approach</td>
<td>28</td>
<td>Yes</td>
<td>2 seconds</td>
<td>97.2</td>
</tr>
<tr>
<td>Fantasia</td>
<td>Zhang et al. [86]</td>
<td>40</td>
<td>Yes</td>
<td>2 seconds</td>
<td>97.2</td>
</tr>
<tr>
<td></td>
<td>Tantawi et al. [124]</td>
<td>40</td>
<td>No</td>
<td>——</td>
<td>94.5</td>
</tr>
<tr>
<td></td>
<td>Proposed Approach</td>
<td>40</td>
<td>Yes</td>
<td>2 seconds</td>
<td>94.5</td>
</tr>
<tr>
<td>AFDB</td>
<td>Zhang et al. [86]</td>
<td>23</td>
<td>Yes</td>
<td>2 seconds</td>
<td>93.9</td>
</tr>
<tr>
<td></td>
<td>Proposed Approach</td>
<td>23</td>
<td>Yes</td>
<td>2 seconds</td>
<td>94.4</td>
</tr>
<tr>
<td>VFDB</td>
<td>Zhang et al. [86]</td>
<td>22</td>
<td>Yes</td>
<td>2 seconds</td>
<td>86.6</td>
</tr>
<tr>
<td></td>
<td>Proposed Approach</td>
<td>22</td>
<td>Yes</td>
<td>2 seconds</td>
<td>85.2</td>
</tr>
</tbody>
</table>

As for the MITDB, the original dataset has 47 subjects. Although studying both the normal and arrhythmic subjects may result in an accuracy deterioration, the proposed approach achieved a high accuracy of 96.5% that is comparable to those approaches that used a subset of the database (i.e. certain number of subjects) and higher than the approaches that used the only the healthy subjects.

In the NSRDB, the original cohort involved 18 subjects. Camara et al. [74] used a KNN classifier with K=1, which make it prone to noise. On the other hand, Dar et al. [87] reported a maximum accuracy of 100%, but no information about the validation was not reported. Moreover, Dar et al. [87] and Fatemian et al. [77] utilized a beat segmentation based on fiducial points (i.e. no blind segmentation). However, the proposed approach achieved a maximum identification rate reached 99%.

As for AFDB, the proposed approach achieved higher accuracy with an average identification rate of 94.4% compared to 93.9% in [86]. Whereas for the VFDB, our approach achieved less identification rate. However, the rank-2 identification rate for the VFDB using the proposed approach reached
The average performance of the proposed approach compared to the state of art work is reported in table 7.3. Although, Agrafioti et al. [37] reported a better overall performance, their evaluation databases involved only health subjects. However, it is not fair to fully compare the work since some papers only reported the maximum accuracy or mentioned no sufficient information about the cross-validation, as well there were no clear information if there was a manually selection of the training and testing samples or not.

Table 7.3: Comparison of the average performance of proposed approach with the literature work.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Blind Segmentation</th>
<th>Databases (count) Type</th>
<th>Identification Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yue et al. [123]</td>
<td>No</td>
<td>(3) N/AbN</td>
<td>85.1</td>
</tr>
<tr>
<td>Dar et al. [87]</td>
<td>No</td>
<td>(3) N</td>
<td>93.2</td>
</tr>
<tr>
<td>Agrafioti et al. [37]</td>
<td>Yes</td>
<td>(3) N</td>
<td>96.2</td>
</tr>
<tr>
<td>Tantawi et al. [124]</td>
<td>No</td>
<td>(3) N</td>
<td>95.9</td>
</tr>
<tr>
<td>Zhang et al. [86]</td>
<td>Yes</td>
<td>(8) N/AbN</td>
<td>93.5</td>
</tr>
<tr>
<td>Proposed Approach</td>
<td>Yes</td>
<td>(7) N/AbN</td>
<td>95.1</td>
</tr>
</tbody>
</table>

N=Normal, AbN=Abnormal

7.4 Summary

In this chapter, we proposed an ECG-based biometric approach that takes into account the cyclostationary property of the ECG signal. In this approach we removed the signal processing and fiducial point detection blocks and employ a blind segmentation algorithm. A deep convolutional neural network was utilized as a classifier to identify subjects. Seven ECG databases as well as cross-validation were performed for model validation. An average identification rate of 95.1% was achieved across all the databases.
Chapter 8

Discussion and Conclusions

The electrocardiogram (ECG) is a bioelectrical signal that is generated from the heart. This signal represents the electrical activity of the cardiac cycle and holds information about any abnormalities in the electrical events. A set of standardized lead configurations can be utilized to measure these signals. However, the number of leads should be suitable for the application of interest.

Since this signal possesses rich information about the heart condition, it can be used to detect any cardiac anomalies or diseases such as ventricular fibrillation, atrial fibrillation, or bundle blocks. Therefore, we utilized the ECG signal recorded from patients to classify four types of heartbeats (right and left branch bundle block, paced, and normal beats) using the signal texture in both temporal and spectro-temporal domains. The utilization of both leads as well as the support vector machine had the highest accuracy of 99.87%. This is because the normal heartbeats are distinguishable in the first lead, while the abnormal beats are more recognizable in the chest lead.

Although the ECG signal has been used for clinical application for decades, researches have started using it in individual identification or biometric systems. We proposed two different ECG-based biometric approaches.

In the first biometric approach, we utilized the spectral images of the ECG heartbeats and 2D convolutional neural networks to extract the high-level individuals’ features. The short-time Fourier transform outperformed the Morse wavelet transform features achieving 97.85% identification rate. This is due to the fact that the Morse wavelet has a better ability to discriminate abnormal heartbeats, and consequently raise the intra-subject variance. When we used this approach in individual verification, and average EER and AUC of 0.0268 and 0.993 respectively were achieved. This proofs that the spectral images have the ability to distinguish individuals when combined with
In the second biometric approach, we employed the cyclostationary properties of the ECG signal to study the changes of the autocorrelation with time. This allowed us to discard the noise removal step, which in turn results in faster algorithm. Moreover, we utilize the blind segmentation which makes the algorithm more robust to noise. An average identification rate of 94.8% over all the databases was achieved, as well as FAR and FRR of 0.04 and 0.002 respectively.

The promising results that we achieved in the proposed algorithms proves that by integrating the frequency information in the feature space, the accuracy would be boosted. Furthermore, if we considered more ECG leads in the classification, we integrate more information, and accordingly, higher accuracy. However, this is not practical in biometric applications to provide more practical data collection during the authentication.

The future work of this research includes the investigation of more features that hold more information about the ECG signal that can boost the performance.
Bibliography


