Towards an Accurate ECG Biometric Authentication System with Low Acquisition Time

by

Juan Sebastian Arteaga Falconi

Thesis submitted in partial fulfillment of the requirements for the

DOCTORATE IN PHILOSOPHY

degree in Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Electrical Engineering and Computer Science
Faculty of Engineering
University of Ottawa

Abstract

Biometrics is the study of physical or behavioral traits that establishes the identity of a person. Forensics, physical security and cyber security are some of the main fields that use biometrics. Unlike traditional authentication systems—such as password based—biometrics cannot be lost, forgotten or shared. This is possible because biometrics establishes the identity of a person based on a physiological/behavioural characteristic rather than what the person possess or remembers. Biometrics has two modes of operation: identification and authentication. Identification finds the identity of a person among a group of persons. Authentication determines if the claimed identity of a person is truthful.

Biometric person authentication is an alternative to passwords or graphical patterns. It prevents shoulder surfing attacks, i.e., people watching from a short distance. Nevertheless, biometric traits of conventional authentication techniques like fingerprints, face—and to some extend iris—are easy to capture and duplicate. This denotes a security risk for modern and future applications such as digital twins, where an attacker can copy and duplicate a biometric trait in order to spoof a biometric system. Researchers have proposed ECG as biometric authentication to solve this problem. ECG authentication conceals the biometric traits and reduces the risk of an attack by duplication of the biometric trait. However, current ECG authentication solutions require 10 or more seconds of an ECG signal in order to have accurate results. The accuracy is directly proportional to the ECG signal time-length for authentication. This is inconvenient to implement ECG authentication in an end-user product because a user cannot wait 10 or more seconds to gain access in a secure manner to their device.

This thesis addresses the problem of spoofing by proposing an accurate and secure ECG biometric authentication system with relatively short ECG signal length for authentication. The system consists of an ECG acquisition from lead I (two electrodes), signal processing approaches for filtration and R-peak detection, a feature extractor and an authentication process. To evaluate this system, we developed a method to calculate the Equal Error Rate—EER—with non-normal distributed data.

In the authentication process, we propose an approach based on Support Vector Machine—SVM—and achieve 4.5% EER with 4 seconds of ECG signal length for authentication. This approach opens the door for a deeper understanding of the signal and hence we enhanced it by applying a hybrid approach of Convolutional Neural Networks—CNN—combined with SVM. The purpose of this hybrid approach is to improve accuracy by automatically detect and extract features with Deep Learning—in this case CNN—and then take the output into a one-class SVM classifier—Authentication; which proved to outperform accuracy for one-class ECG classification. This hybrid approach reduces the EER to 2.84% with 4 seconds of ECG signal length for authentication.

Furthermore, we investigated the combination of two different biometrics techniques and we improved the accuracy to 0.46% EER, while maintaining a short ECG signal length for authentication of 4 seconds. We fuse Fingerprint with ECG at the decision level. Decision level fusion requires information that is available from any biometric technique. Fusion at different levels—such as feature level fusion—requires information about features that are incompatible or hidden. Fingerprint minutiae are composed of information that differs from ECG peaks and valleys. Therefore fusion at the feature level is not possible unless the fusion algorithm provides a compatible conversion scheme. Proprietary biometric hardware does not provide information about the features or the algorithms; therefore, features are hidden and not accessible for feature level fusion; however, the result is always available for a decision level fusion.

Acknowledgements

I want to thank my supervisor Dr. Abdulmotaleb El Saddik, for his continuous support in this challenging and amusing journey. His positive energy was always a vital factor during the mood fluctuations of research. It is a privilege to work with a professor with vast knowledge in many areas, including life. The last one is very important for our personal growth, thank you.

I also want to thank my colleagues at the MCRLab, it is an honor to be surrounded and learn every day from bright people like you.

A special thanks to my mother Miriam, my father Joffre and my siblings Diana and Ana María for their constant love and support. I am glad to have you in my life. Your love breaks distance barriers and keeps us always together.

I want to acknowledge SENESCYT for their financial support under the scholarship program "Convocatoria Abierta".

To my beloved family Walaa, Elías and Yara,

It is your love that makes my heart unique.

Table of Contents

List of Tables

List of Figures

Glossary of Terms

CNN: Convolutional Neural Network
CWT: Continuous Wavelet Transform
DET: Detection Error Trade-off
DTwin: Digital Twin
DWT: Discrete Wavelet Transform
ECG: Electrocardiogram
EER: Equal Error Rate
FAR: False Acceptance Rate
FM: False Match
FMR: False Match Rate
FNM: False Non-Match
FNMR: False Non-Match Rate
FRR: False Rejection Rate
GAR: Genuine Acceptance Rate
HCI: Human Computer Interaction
IoT: Internet of Things
ISD: Inverted Second Derivative
MRA: Multi-resolution Analysis
NBIS: NIST Biometric Image Software
NIST: National Institute of Standards and Technology
QRS: Q valley, R peak and S Valley from an ECG signal
RBF: Radial Basis Function (Gaussian Kernel)
ReLU: Rectified Linear Unit
ResNet: Residual Neural Networks
RMS: Root Mean Square
ROC: Receiver Operating Characteristic
STFT: Short Time Fourier Transform
SVM: Support Vector Machine
TAR: True Acceptance Rate
TRR: True Rejection Rate

Chapter 1.

Introduction

1.1 Background

Technological advancements have increased the interaction of humans with electronic devices and this will continue to grow. El Saddik [1] has envisioned the concept of Digital Twin as the digital replication of living and non-living entities. This digital replication—and other interactions—necessitate users to interact regularly with an increasing number of devices. Hence, they have to maintain a memory-taxing amount of usernames and passwords. Biometric schemes offer an optimum alternative for the use of usernames and passwords for authentication [2]. In addition, they do not suffer from the vulnerabilities of conventional password protected systems. For instance, pin numbers and graphic patterns are popular authentication techniques that are vulnerable to shoulder surfing attacks (i.e. people watching from a short distance) [3]. In 2012, a Visual Privacy Productivity Study sponsored by the 3M Company found that 82% of IT professionals believe that employees are careless about shoulder surfing attacks [4]. The same study revealed that 72% of commuters in the UK have been able to observe passwords of commuters through shoulder surfing.

Biometrics eliminates the threat of shoulder surfing. However, this approach presents other important vulnerabilities. Damage on biometric traits can lead to an authentication failure [5] or attackers can duplicate biometric traits to gain access [6]–[9]. Pictures and videos can easily spoof facial biometrics, high resolution pictures of an Iris with a whole in the pupil space can spoof Iris scanners, latex and conductive ink can spoof fingerprints [9].

The field of medicine uses Electrocardiograms (ECG) to diagnose health issues related to the heart. An ECG signal differs from each individual and physicians approximate ECG signals to a common patron in order to diagnose a heart condition. What appears to be a disadvantage in medicine, it is an advantage in the field of biometrics. Heart rate changes affect the ECG by expanding or contracting the signal in the time domain. These changes do not affect the unique biometric characteristics of the ECG and a normalization procedure corrects these changes [10], [11]. The uniqueness of ECG makes it a suitable biometric trait to differentiate individuals. ECG authentication conceals the biometric trait, which prevents duplication. This is an advantage over more accurate biometric approaches like Iris, face and fingerprints. While some of these biometrics has more than 100 years of research [12] and has the highest accuracy among other biometrics [5]; ECG is less prone to be attacked by biometric trait duplication. A subject can leave trails of fingerprints on every object that they touch; however, we can extract ECG only with an electrocardiograph. Therefore, subjects cannot leave traces of ECG; they must be present and alive—aliveness detector is an additional security—in order to extract the ECG biometric trait. However, while collection of other biometric traits can be achieved almost instantaneously, the ECG method typically requires 10 seconds or longer to capture ECG signals and achieve an acceptable level of accuracy for authentication [10], [13]–[16].

In a previous work [11], we proposed an ECG authentication algorithm that requires a 4 seconds long signal to achieve a False Acceptance Rate (FAR) of 1.41% and a True Acceptance Rate (TAR) of 81.82%. This algorithm uses a manual threshold tuning of the ECG biometric features.

An alternative, to improve accuracy and keep a low authentication time in ECG biometric, is to use a procedure that automatically sets thresholds. In this manner, we can train a

machine-learning algorithm to determine the appropriate thresholds and use them for authentication.

Another alternative to achieve the goal of higher accuracy is to use different features. Normally, the feature extraction stage requires a feature engineering process. Deep learning techniques implements automated processes for feature extraction that replaces the manual feature engineering. One of the well-known deep learning techniques is Convolutional Neural Networks—CNN—and is designed for image classification—currently reporting excellent results. A CNN deep learning technique could be used in order to improve ECG authentication accuracy and time acquisition; however, two problems arise: ECG is not an image and deep learning does not work for one class classification problem—authentication is a one-to-many problem. Some literature [17]–[20] claims to perform authentication with deep learning, but in reality is identification.

To solve these two problems, ECG authentication can use a hybrid solution that combines CNN and SVM. CNN can extract ECG features and a Support Vector Machine—SVM—can perform one-class classification—authentication. The original purpose of CNN was image classification [21]. There are numerous CNN pre-trained models where millions of images have train them, there are available and deliver excellent results [22]. Many studies use other type of signals to train a CNN model; however, the amount of data to train their models does not compare with the data of the pre-trained CNN models for images. GoogLeNet [23] is a pre-trained CNN model that is in the limit between accuracy and processing complexity [22]. Other models are more accurate but more complex to process and other models are less complex to process but less accurate.

A multibiometric approach can be another solution to improve accuracy while maintaining a low acquisition time with ECG. Multibiometric systems combines one or more biometric traits (e.g. Fingerprint, face, gate, signature, voice and ECG among others) in order to complement weaknesses and enhance robustness in one secure solution that achieves the authentication of an individual [24]. Multibiometrics combines biometric traits employing several approaches that are categorized as: Multi-sensor, Multi-algorithm, Multi-instance, Multi-sample and Multi-modal. Multi-sensor combines two or more different type of sensors

to capture the same trait (e.g. Face authentication with 2d and 3d cameras). Multi-algorithm combines two or more different algorithms on the same biometric trait (e.g. Face authentication with eigenfaces and fisherfaces). Multi-instance combines, not the same trait, but the same type of biometric trait (e.g. Fingerprint authentication with Left and Right index fingers). Multi-sample combines the same biometric trait from different perspectives (e.g. Face authentication with front and side images). Multi-modal combines different biometric traits (e.g. Fingerprint, Face and ECG).

ECG can fuse with fingerprint—the most accurate biometric [25]—to enhance the strengths and reduce the impact of the weaknesses of each approach while increasing the accuracy with low acquisition time of the biometric authentication process. This will prevent attackers from faking fingerprints and would allow users to authenticate securely within 4 seconds.

Fusion of ECG and fingerprint refers to a multimodal biometric system, to be more specific: a bi-modal biometric. A bi-modal system enhances a uni-modal biometric (single biometric trait) system, reduces the probability of encountering two users with the same biometric information (non-universality characteristic) and can speed up indexing for identification in large-scale databases. The use of multiple traits can narrow down the amount of potential matching candidates and focus the search only among a few stored templates. In addition, bi-modal biometrics provides robustness under noisy environments; if one trait fails, we can use another trait without disrupting the biometric process [24]. Bi-modal biometrics has a wide range of applications; as to name few: mobile devices, facility access, automotive security, forensics, active authentication—continuous—and communication with digital twins [26], [27].

1.2 Motivation

Modern gadgets implement biometric authentication—fingerprints, face, iris, voice among others—to protect the privacy information of users. However, these biometric technologies expose the biometric traits. An attacker can exploit this vulnerability and collect these exposed traits, duplicate them and gain unauthorized access [6]–[8]. Aliveness detection is the

more accepted approach to solve the spoofing problem of biometrics [28]. Many biometric traits use ECG signals to detect if user is alive [28] and then perform authentication. This is not a solution where the biometric trait solves the problem. It is another signal that solves the spoofing problem. ECG can be a biometric trait and intrinsically solves the spoofing problem by been ECG an aliveness signal.

ECG is an electrical signal that represents the activity of the heart. Location, size, anatomy, chest structure and other factors makes the ECG signal unique among individuals. ECG biometrics conceals the biometric trait and prevent users to leave traces of biometric traits in places where an attacker can effortlessly collect them and duplicate them. ECG requires special equipment—electrocardiographs—to extract the ECG biometric traits. Simultaneously, ECG is an aliveness indicator; a user must be alive in order to extract the biometric trait. This is not the case for biometric technologies like fingerprints, face or voice [6]. In addition, a person cannot avoid detection by intentionally damage the biometric trait. If a person tries to damage its ECG, it will represent a dangerous health issue.

Despite the security advantages of ECG, the relative novelty of the technology faces some challenges. ECG authentication is not as accurate and fast as other biometrics—e.g. Fingerprints. ECG requires several heartbeats—several seconds of ECG signal—to get an acceptable accuracy. Most of the current research [10], [13], [16], [29]–[32] on ECG authentication focuses on improvements on the accuracy rates. Only few studies [11], [14], [33] address the authentication time issue.

Authentication time is an important aspect for the implementation of the technology in an end-user product. It would be uncomfortable for users to wait ten seconds or more to gain access to their devices. On the other hand, shorter authentication times leads to less accuracy, which will also make users uncomfortable.

An ECG biometric authentication system should be fast and accurate—at the same time—in order to provide better protection than traditionally authentication systems. ECG presents stronger security advantages over traditional authentication systems but it has to overcome the challenges mentioned above.

1.3 Objective

We would like to explore ECG authentication algorithms that achieve high accuracy with short authentication time. To do this, we will develop ECG authentication algorithms that use SVM, CNN and biometric fusion (Decision Level Fusion). In addition, we developed a signal-processing tool and an evaluation method. Although these tools are not authentication algorithms, they are a fundamental part of the ECG authentication process. They provide a better input in order to achieve the general goal of high accuracy with short authentication time.

1.4 Contributions

In this Thesis, we will design, develop, and validate ECG as biometric authentication system with higher accuracy and short signal acquisition time. In order to achieve our goal we have the following contributions:

- Design and development of an ECG Authentication algorithm that sets an automatic threshold in order to improve accuracy with a short signal acquisition time. This approach uses SVM to automatically set a threshold in order to improve the accuracy of manual threshold tuning.

- Design and development of an ECG authentication algorithm that use automatic feature extraction to enhance accuracy and keep a short signal acquisition time. This algorithm uses deep learning—CNN—to automatically detect ECG biometric features in order to perform authentication. This is a hybrid approach that uses SVM and CNN. A CNN—deep learning—determines the ECG features and the authentication use these features with SVM to automatically set thresholds and perform the authentication.

- Design and development of an algorithm that combines two biometric modals in order to have a better accuracy with a short acquisition time. This algorithm uses fingerprint in combination with ECG. This bi-modal algorithm uses the speed and accuracy advantages of Fingerprint biometric to improve the accuracy and

authentication time of ECG. At the same time, it keeps the security advantages of ECG biometrics.

As part of this study, we developed a signal-processing tool and a calculation method to help these algorithms to achieve the general goal of high accuracy with short authentication time. The contributions of these works are:

- Design and development of an R-peak detection algorithm with a low average time error. This is a signal-processing tool to detect the location of the R peak in an ECG signal. This tool detects the R peaks of an ECG signal with the lowest error in the time location of the R peak. Less error in the time location of the R peaks, contributes to improve the accuracy of the algorithms.

- Development of a method to calculate the Equal Error Rate—EER—and Detection Error Trade-off—DET—in a multi-threshold approach. This method provides an evaluation solution of algorithms that uses multi-thresholds to perform authentication. Multi-thresholds generate several operation points that we cannot directly compare with single thresholds operation points. This method generates single threshold operation points from multi-thresholds operation points. In order to compare the accuracy among approaches with one threshold and multiple thresholds, it is important to have a calculation method that provides an evaluation parameter—DET and EER—that is common within these two approaches.

1.5 Scholarly Achievements

This thesis has generated publications on peer reviewed journals and conferences that validates our work in the research community. The following is a list of accepted—or in progress—publications.

Papers at refereed Journals

1) A. El Saddik, H. F. Badawi, R. Velazquez, F. Laamarti, R. Gámez Diaz, N. Bagaria, and **J. S. Arteaga-Falconi**, "Dtwins: A Digital Twins Ecosystem for Health and Well-Being," IEEE COMSOC MMTC Commun. - Front., vol. 14, no. 2, pp. 39–43, 2019.

2) Y. Qiu, Y. Liu, **J. Arteaga-Falconi**, H. Dong and A. E. Saddik, "EVM-CNN: Real-Time Contactless Heart Rate Estimation From Facial Video," in IEEE Transactions on Multimedia, vol. 21, no. 7, pp. 1778-1787, July 2019.

3) **J. S. Arteaga-Falconi**, H. Al Osman, and A. El Saddik, "ECG and Fingerprint Bimodal Authentication," Sustain. Cities Soc., vol. 40, pp. 274–283, 2018.

4) B. Hafidh, H. Al Osman, **J. S. Arteaga-Falconi**, H. Dong, and A. El Saddik, "SITE: The Simple Internet of Things Enabler for Smart Homes," IEEE Access, vol. 5, pp. 2034–2049, 2017.

5) **J. S. Arteaga-Falconi**, H. Al Osman, and A. El Saddik, "ECG Authentication for Mobile Devices," IEEE Trans. Instrum. Meas., vol. 65, no. 3, pp. 591–600, 2016.

Papers at refereed Conferences

1) **J. S. Arteaga-Falconi**, H. Al Osman, and A. El Saddik, "R-peak detection algorithm based on differentiation," in Intelligent Signal Processing (WISP), 2015 IEEE 9th International Symposium on, 2015, pp. 1–4.

2) **J. Arteaga-Falconi**, D. P. Tobón, and A. E. Saddik, "EER Calculation and DET Approximation in a Multi-Threshold Biometric System," in 2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 2018, pp. 1–6.

Patents

1) A. El Saddik, **J. S. Arteaga-Falconi**, and H. Al Osman, "Electrocardiogram (ECG) biometric authentication," United States Patent US 9,699,182, United States Patent and Trademark Office. 4 Jul. 2017.

1.6 Thesis Organization

We organized the thesis as follows:

- Chapter 2 provides the background and related works that we used to build this thesis and to validate our results.

- Chapter 3 presents the design and evaluation of ECG authentication algorithm with SVM that sets automatic thresholds.

- Chapter 4 delivers the design and evaluation of an ECG authentication algorithm with deep learning that automatically extracts features.

- Chapter 5 presents the design and evaluation of a bi-modal biometric algorithm with ECG and Fingerprint

- Chapter 6 provides the design and evaluation of an R-peak detector algorithm that uses differentiation for a low average time error.

- Chapter 7 provides and evaluates a detection error trade-off and Equal Error Rate calculation method for algorithms with multiple thresholds.

- Chapter 8 presents the conclusions of this thesis and provides directions for future works.

Chapter 2.

Background and Related Works

The Encyclopedia of Biometrics defines Biometrics as “the science of establishing the identity of a person based on the physical or behavioral attributes associated with an individual” [34]. A biometric system works with specific modules, has different operation modes and uses physiological and behavioural characteristics known as biometric traits. Biometrics can combine two or more biometric inputs in a multibiometric system.

This chapter presents a general overview of Biometrics and deep learning as part of machine learning, which are the groundwork for this research. It describes the operations modes and the different traits that a biometric system uses. Also describes the evaluation steps that measure accuracy and the concepts of multibiometrics.

2.1 Biometrics Modules

Figure 1 shows the four principal modules in a conventional biometric system. First, the sensor module acquires the biometric trait, such as fingerprint, ECG, face. Then, the feature extractor module processes the biometric trait and extracts features to generate a biometric template. A biometric template is a collection of features that represents a biometric trait. If the

biometric system is enrolling (registering) users, it stores the template in memory (e.g. database). Otherwise, it sends the template to the matcher module. The matcher module typically generates a score that represents the similarity between the input and a stored biometric template(s). The decision maker module uses these score(s) to determine the result of the process [35].

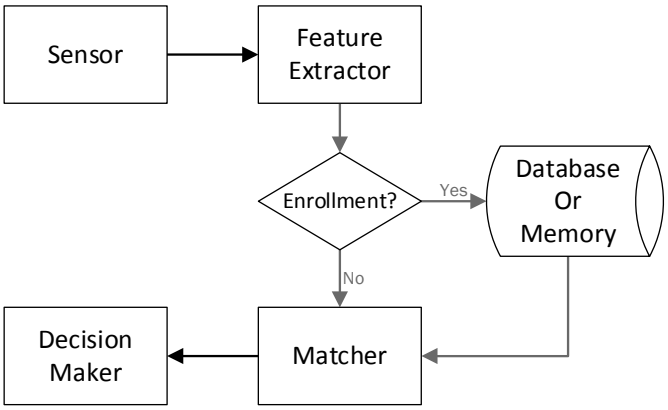

Figure 1. Modules of a Biometric System

2.2 Biometrics Modes of Operation

Biometric systems has two modes of operation: identification and verification [36]; verification is a synonym of authentication [37]. This section describes these operation modes. It is important to mention that this work focus in authentication mode.

2.2.1 Identification

During identification, the biometric system searches for a person within a database. The user supplies a biometric trait, generates the template and the system compares it against a database of user templates. If there is a match, then the system is said to have identified the user. Identification follows a one-to-many scheme, where the system compares one user input against all the user records in the database [34].

2.2.2 Authentication

For authentication, the system stores biometric information that corresponds to a genuine user. When a user supplies a biometric trait, the system compares it against the

stored biometric information and determines if the user is genuine or an impostor. Authentication follows a one-to-one scheme whereby the system compares one user input against the genuine user record in the system [34].

2.3 Biometrics Evaluation

A password-based system always has a perfect match or no match result. In a biometric system, aspects like moisture, weather, sensors, physical conditions, and others; produce samples that are similar but never the same. This is why the implementation of a threshold is required in a biometric system [35]. When a biometric system is not perfect, there are parameters that help to measure the accuracy and the errors of the biometric system. The following sub-sections describe the evaluation rates and the curves that measure the accuracy a biometric system.

2.3.1 Evaluation Rates

Thresholds will cause the system to have errors. A very permissive threshold will result in a system that will validate the genuine users, but also will validate the impostors. This is known as False Acceptance [38] and the rate is abbreviated as FAR (False Acceptance Rate). Equation (1) calculates the False Acceptance Rate.

$$FAR = \frac{\text{Total false acceptances}}{\text{Total impostor attempts}} \quad (1)$$

The opposite of the false acceptance is the false rejection (FR) or the false non-match (FNM), which refer to scenarios when the system has a very restrictive threshold. The consequence of this is a high rejection of genuine users. This is known as FRR [38] (False Rejection Rate) and it is represented in (2).

$$FRR = \frac{\text{Total false rejection}}{\text{Total genuine attempts}} \quad (2)$$

FAR and FRR are terms that relates directly with the system threshold. With a low threshold, we will have a low FRR and a high FAR. Consequently, the biometric system

will have a very low probability to reject genuine users and will have a high probability to accept impostors. On the other hand, the opposite will happen if the threshold is too high; the biometric system will have a low FAR and a high FRR. As a result, the biometric system will reduce the probability of an impostor to access and will increase the probability of rejecting genuine users. As depicted in Figure 2., when FRR and FAR are evaluated, it is desirable to have these two values at their lowest in order to ensure an accurate system. The point with the lowest equidistant values is known as Equal Error Rate (EER) and can be visualized in the Detection Error Trade-off graph (DET) [39]. Section 2.3.2 describes the DET graph

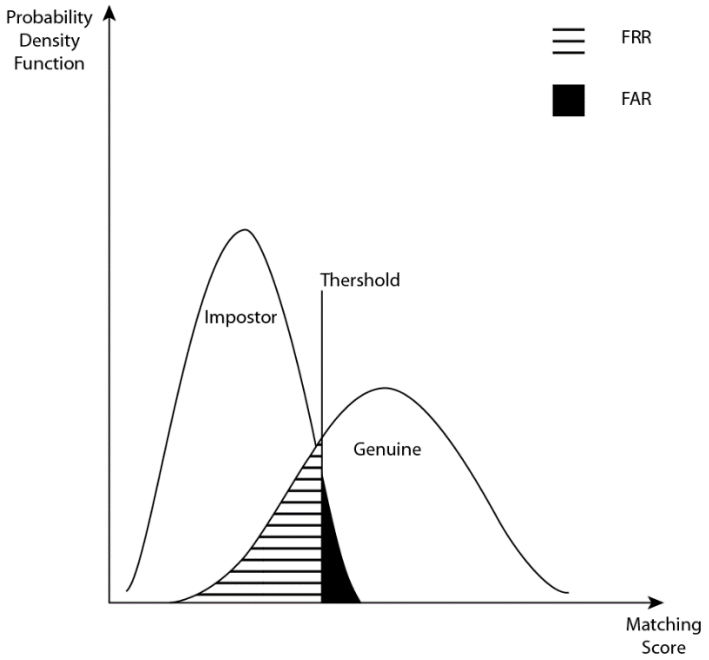

Figure 2. Acceptance/Rejection of Impostor/Genuine in function of threshold

True Rejection Rate (TRR) relates with FAR because TRR is the total number of true rejections against the total number of impostor attempts: TRR = 1-FAR. True Acceptance Rate (TAR) is the total number of true acceptances against the total number of genuine attempts, which is related with FRR as: TAR=1-FRR [40]. Given this analogy, a system can also be evaluated based on its FAR and TAR, where a very low FAR with a very high TAR is desirable for a reliable system. Some works name TAR as GAR (Genuine Acceptance Rate).

2.3.2 Evaluation Curves

Detection Error Trade-off (DET) is the representation of a threshold that moves around different values in order to calibrate the False Rejection Rate in terms of the False Acceptance Rate. This is not limited to one threshold only, when more than one threshold is involved, it generates more data and the plot of the DET is with a different approach. Biometrics uses the DET graph in order to evaluate performances. Generally, unibiometrics uses one threshold and multibiometrics uses more than one threshold. The use of multiple thresholds improves the performance of a biometric system [41].

Classification task has different operating points (i.e., False Acceptance, True Acceptance) depending on the requirements of the application. For instance, access to a bank vault requires very low tolerance to false acceptances, meaning that the probability of an impostor to be accepted is very low. The cost of having a low false acceptance is that more users that are genuine will be rejected (higher false rejection). However, for banks, that is an acceptable scenario. If we want to apply the same conditions in mobile biometrics, the situation is different. A user access his device more often, therefore a higher false rejection is a problem. This can be reduced at the cost of increasing the false acceptance (increasing the probability of impostors to have access). The operating point depends on the application, that is why performance cannot be measured by a single number and a performance curve is preferred [42]. Some of the known performance curves are ROC (receiver-operating characteristics) and DET (Detection Error Trade-off).

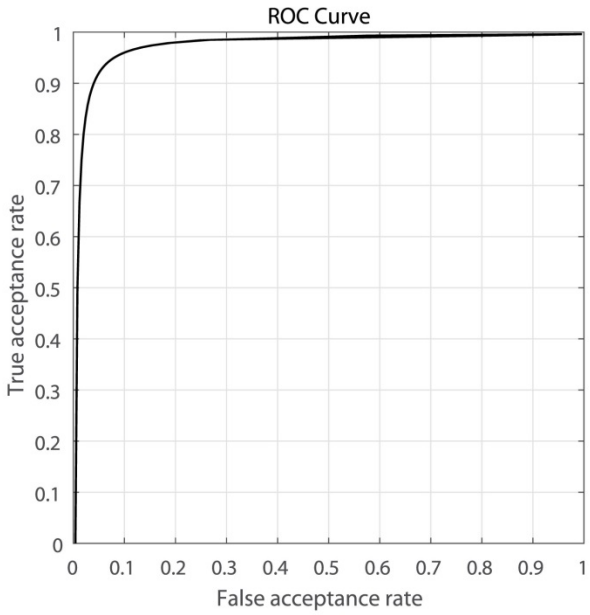

Figure 3. Receiver-operating characteristics curve (ROC)

ROC curve is a graph that helps visualize the trade-off between false positives and true positive tests [43]. In biometrics, this is the discrimination between false acceptance and true acceptance rates. See Figure 3. The graph visualizes how the system discriminates genuine users against impostors. It helps to see the trade-off to have correctly validated genuine users against impostors that are incorrectly recognized as genuine users. DET graph is another approach to evaluate the performance of a classification system. In this case, data is presented in terms of FAR (False Acceptance rate) and FRR (False Rejection Rate). See Figure 4. In another words, it represents the performance of the system in terms of errors. It provides information about the number of genuine users that are incorrectly rejected (FRR) against the number of imposters that are incorrectly accepted as genuine.

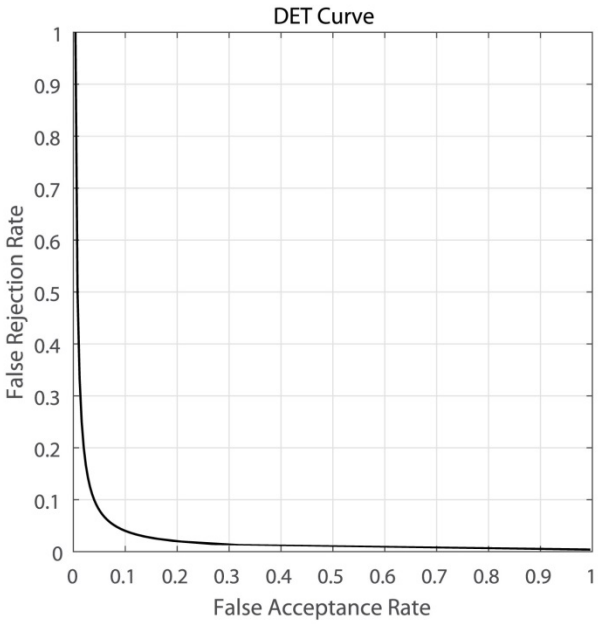

Figure 4. Detection Error Trade-off Curve (DET)

The use of ROC and DET will vary depending on how we want to evaluate the system. Many studies use DET in order to calculate EER (Equal Error Rate), which is a parameter to compare among several systems. EER is a parameter that indicates the optimum operation value (accepted impostor users is equal to rejected genuine users), where a lower value indicates a better system. ROC and DET curves provide more information regarding the performance of the biometric system. The fingerprint verification competition 2004 (FVC2004) evaluates fingerprints using ROC and DET graphs and calculates the EER for several algorithms that were part of the competition [44]. Many studies simplify the presentation of results by just showing the EER. It is a valid approach and the majority of studies use it; therefore, the present work will use DET graph in order to calculate EER.

2.4 Biometric Traits

Biometrics establishes the identity of a person based on physiological and behavioural traits. Biometrics traits such as eyes, iris, fingerprint or ECG are physical biometric attributes. In contrast, signature, gait or keystroke patterns are behavioral biometric attributes.

Biometric authentication uses a variety of traits; the following subsection describes the most common traits in the biometrics authentication research.

2.4.1 Behavioural Biometrics

Behavioural biometrics uses unique characteristics that we might be conscious or not when accomplishing everyday tasks. Behavioural biometrics is divided in five categories [45]. The first category corresponds to identify the person by the writing styles such as vocabulary and punctuation or the style when they draw. The second category corresponds to the Human Computer Interaction (HCI) and analyses the manner a person interacts with a computer (e.g. how to open programs, shortcuts used with the keyboard). The third category is the physical observation of the interaction with the computer. E.g. posture, mouse holding, look at the display. The fourth category uses motor skills, analyses the muscles involved in order to accomplish a physical task. The last category involves analysing the manners we perform a physical task like the technique we use to walk, the technique to grab a pencil, the way we type a keyboard (force, fingers and patterns). In this section we describe the most common used behavioural biometrics.

Gait authentication is implemented using onboard accelerometer data similar to the accelerometers from mobiles [46]. Derawi et al. [46] suggest that validation rates are promising but are not reliable enough for practical use yet and still need further improvement to produce reliable results [20]. According to Clark et al. [3] Gesture-Based authentication traits, like gait, can be stolen with a properly trained attacker; therefore, attempts should be limited.

Keystroke biometrics measures the unique pattern that human have when typing a keyboard, measuring speed and force. The advantage is that it can be used for continuous authentication. While a person is working using the keyboard, the system can continuously authenticate the user in order to validate that is a genuine user. The disadvantage of this biometric is that it can be affected by changes on emotion of the user, which might lead to changes on speed and force which will affect the accuracy of the system [47].

Signature Biometrics authenticates a person by detecting the writing pattern [48]. The traditional method to authenticate documents has used signatures on paper in order to legalize documents. Biometrics with signatures uses the same concept, but instead to

signing on a paper, it signs in/with an electronic signature device or digitalizes the signature on paper (creates an image) in order to analyze it with a computer. When the user signs with an electronic haptic device, it can use haptic features in order to determine the signature and not just by the image. It can detect the pressure and movements in the pen [49]. A constraint that image signature biometric has is that can be easily forged by an impostor [50]. Haptics signature biometrics solves this issue. The constraint of using haptics for authentication is that conventional biometric systems (e.g. iris, voice, fingerprint/hand geometry) perform better. However, haptic biometrics is an optimum solution for active authentication solutions [51].

2.4.2 Physiological Biometrics

Physiological biometric uses traits that are unique among humans. The advantage is that extracting the features is faster than behavioural biometrics. We will describe the most common physiological biometrics traits in this section.

Face biometrics uses face characteristics like distance between eyes, nose, mouth, eye brows chins or spectral analysis of the full face [52]. The advantage is that it presents an easy approach for humans to validate the authenticity. Face biometrics extracts the traits from an image. A digital camera provides the image and the characteristics will depend on the requirements of the system. In general, they are easily available which facilitates the implementation. The disadvantage is that is affected by the position of face when the image is capture, lightning conditions and image background noise.

Ear biometrics has a similar approach as face biometrics. It uses an image in order to measure unique characteristics in the ear. The advantage is that is easy to access the sensor that captures the trait and does not greatly vary with time. The disadvantage is the same with face recognition. Lightning for the image, capturing position and background noises [53].

Palm Prints Biometrics is similar to fingerprints; it uses the skin ridges of the palm. This type of biometric systems are more accurate than fingerprint because they use a larger area

to extract the trait [54]. The disadvantage is that sensors will be bigger and not easy to implement in portable devices.

Hand Geometry authentication uses the hand shape as a biometric trait. It measures size, length of the palm and also the fingers [55]. The advantage of the technology is the simplicity to use and implement. In addition, it is not affected by skin or environment conditions. The main issue with hand geometry is the uniqueness. There is a high chance that several people will have the same biometric traits within a large population. Growth and diseases like arthritis can affect the accuracy [52]. Sensors for Hand Geometry might require a specific size that might not make it suitable for some implementation, i.e. onmobiles.

Researchers have proposed the use of iris-based authentication [56]. While this method is reliable, it is sensitive to changes in lighting conditions which makes its deployment challenging [57].

Voice is another trait used in biometrics [58]. However, because the voice must be recorded, the primary concerns of this method is the surrounding noise and easy duplication of the biometric features [59].

Fingerprint Biometrics and ECG Biometrics are part of the physiological biometrics. In this thesis, we use both biometrics and we will explain them in section 2.6 and section 2.7.

2.5 Fingerprint Biometrics

The study of fingerprints has a long history, originating in 1686 [12]. Today, the use of fingerprint biometrics has become a well-established technology with an outstanding performance compared to biometrics of other traits [5].

Typical features used in fingerprint biometrics include minutia, local orientation, ridge shape, local frequency, singularity, and ridge count. Ridges and valleys compose fingerprints and a minutia is the end (or bifurcation) of a ridge or a valley. Minutiae are the most used features by fingerprint matchers. Minutiae can carry information about direction and spatial

frequency. The minutia direction is a vector that defines the course of a ridge or valley. Spatial frequency defines the direction of the finger. To determine spatial frequency, algorithms usually uses the minutia direction as input information [60].

A recent study by [61], used a fuzzy logic control system for matching. They used "Atanassov's intuitionistic fuzzy sets" to determine the number of minutiae that are required for a match. They tested their algorithm with 89 users and attained an accuracy of 98.08%. Another study [62] used the Ant Colony Optimization (ACO) scheme to find correspondence between minutiae. ACO builds a correspondence graph of a promising solution based on the most prominent features. Ants use this graph to find all minutiae correspondences. They test their algorithm with 800 fingerprints, 100 fingers with eight images of each finger and report an Equal Error Rate (EER) of 2.79%. A third study [63] used a genetic algorithm (GA) for matching. They considered the starting population to be the stored template(s) and used the input fingerprint to calculate the fitness. The GA launches an iterative process that produces a new generation from the previous one through selection, crossover and mutation. The GA finishes when it reaches a point where there is no change in the fitness values for a number of generations. If the fitness value is greater than a threshold, there is a match. They tested with 2000 pairs of fingerprints and according to the ROC curve; they obtain a FAR of 0.05% and a TAR of approximately 84%. A fourth study [64] divided the input fingerprint image into squares and each square into two triangles. They adjusted the orientation of the triangles in the stored template(s) and the input to minimize distortion. Then, they proceeded with a minutiae match. They evaluated the algorithm with 2000 pair of images and reported a TAR of 85% with a FAR of 0.05%.

For over a decade, the National Institute of Standards and Technology (NIST) has been working together with the FBI on developing a biometric software application called the National Background Investigation System (NBIS) [65]. They constantly improve and test the software application on a database of over 40,000,000 fingerprints. In 2015, they released the latest stable version 5.0.0, which is open source and publicly available (access is restricted in some countries). The software has independent modules, including the feature extractor and matcher module. The feature extractor module uses minutiae features collected using the

MINDTCT algorithm [65]. The matcher module uses the BOZOSRTH3 algorithm [65]. Maddala *et al*. [66] evaluated the NBIS with 100 fingerprints and 8 images of each fingerprint. According to Detection Error Trade-off (DET) graph, they reported an EER of approximately 0.4%.

While fingerprint authentication typically performs well, hackers can easily duplicate fingerprints left behind unintentionally on smooth surfaces to break into the system. In contrast, ECG methods conceal the biometric features, but require a lengthy signal for authentication. Decreasing the length of the signal affects the performance of the method negatively. Section 2.6 explains these with more detail.

2.6 ECG Biometrics

An ECG displays the electrical activity of the heart. To obtain an ECG record, a sensor collects, amplifies and filters an electrical signal obtained through electrodes in contact with the human body. These electrodes are placed in specific arrangements called virtual vectors or leads [67]. There are 12 lead configurations. Lead-I requires only two electrodes placed on the chest, arms, or hands. ECG is a unique signal for every person due to the variation in heart mass, orientation, gender, conductivity, and order of activation of the cardiac muscles [68], [69]. This characteristic makes ECG biometrics possible [16].

2.6.1 Approaches Using ECG as Biometric

To the best of our knowledge, [13] presented the first approach for ECG biometrics. They extracted time, amplitude and slope based features for classification. They concluded that it is possible to identify a person using ECG and three electrodes are enough for ECG authentication. ECG is affected by heart rate changes and this study shows the linearity of this changes and this is later corroborated by [10]. To correct this affection Israel *et al.* [10] apply a normalization based on a unitary system where the duration of each feature is divided by the total length of a heartbeat. The normalization fits all the heartbeats within the same window size. A previous work [11], implements the normalization procedure with a linear formula where it takes the R-Peak to detect the heart beat and determines the

beginning and the end of the heart beat cycle. Later, an average duration is calculated and all the heartbeats are expanded or contracted in order to fit the calculated average.

In another study, [29] used linear discriminant analysis for filtering before matching, then calculated the Euclidean distance and applied a threshold to determine authenticity. They evaluated the algorithm with an ECG database of 27 subjects and a signal length of 2530 seconds. They reported an EER of 0.6%. Chiu et al [30] applied Haar wavelet transformation to the ECG signal and used the obtained coefficients as ECG features. They calculated the Euclidean distances between the template and the input fingerprint features for matching. They evaluated the algorithm with two ECG datasets from 25 and 45 subjects; each record was 2 minutes long. They reported an EER of 0.83% to 0.86%. Molina et al. [31] extracted features based on the fiducial points. They calculated a score for the features in the stored template and input template. To perform authentication, they applied a threshold to the score. For the evaluation, they used ECG records 5 minutes in length from 10 subjects and obtained an EER of 2%. Lourenço et al. [32] applied the k-nearest neighbours (k-NN) algorithm and an SVM classifier on ECG features to perform identification and authentication. They used 5-minute ECG records from 32 subjects. Authentication with the k-NN algorithm achieved an EER of 9.39% while authentication with SVM scored a FAR of 0% with a false rejection rate (FRR) of 4.55%. They found that SVM is better than k-NN for authentication; however, for identification, they found the k-NN algorithm to be the better method.

Convolutional Neural Network—CNN— is a deep learning technique designed for image classification. However, researches have been using CNN for other purposes, including ECG biometrics. Labati *et al*. [70] perform identification, authentication and continuous authentication with ECG signals. They extract the QRS complexes and create a one-dimensional signal that feeds the CNN. They trained their CNN model with 277,563 samples that is equivalent to 771 hours of ECG record time. They use only healthy patients ECG signals from the PTB database [71] and removed signals from 17 users due to noise contamination. They performed experiments using three different leads and a multi-sample fusion with the three leads and the ECG record for authentication is five minutes long.

Their best reported authentication result with one lead has an Equal Error Rate—EER—of 3.37%.

Another CNN model is used in a ECG multi-algorithm—multibiometric—approach proposed by Da Silva Luz *et al.* [72]. They propose two CNN models to extract features. The first model extract features from a Raw ECG signal and the second model extracts features from the spectrogram of the same ECG signal. They measure the distance of the features from a previously created template and they fuse the results at the score level. They use an outlier algorithm to exclude corrupted heartbeats, leaving them with 17,226 heartbeats from 1 database and 50,401 from another database. Half of those heartbeats are for enrollment—training—and the other half for authentication. They report results on unibiometric and multi-algorithm fusion. When they train the model with different ECG signals than those used for authentication, their results are 26.38% EER with an ECG spectrogram, 14.13% EER with an ECG raw signal and 12.78% EER with multi-algorithm fusion.

Residual neural networks—ResNet— is another approach for ECG Authentication and Identification proposed by Chu *et al.* [73]. They modify the kernel size of ResNet and use three parallel kernels of different size in order to extract ECG Features. Later, they generate feature vectors—each vector has two heartbeats—and perform authentication by calculating the Euclidean distance between those vectors. To perform identification they use the classification layer of their proposed model. They train their model with one ECG database, where, they extract 200,000 vectors that represent 400,000 heartbeats. To test ECG authentication, they use 3 other databases—including PTB Diagnostic Database [71]—and exclude ECG records of non-healthy users. They generate 1000 vectors—equivalent to 2000 heartbeats—from each user and they use half of it to generate the authentication template and the other half to perform authentication. They combine two heartbeats to generate a vector. The average record length per users is two minutes long and the average number of beats within two minutes is approximately 167 heartbeats. Since there is not enough time to complete 1000 vectors with 167 heartbeats per user, they complete the 1000 vectors by using the same heartbeats in different vectors. They end up

with 1000 unique combinations but a heartbeat will be present in two or more vectors. They report an accuracy of 95.04% to 100% for identification and an Equal Error Rate of 0.59% to 4.74% for authentication.

Zhao *et al.* [74] propose ECG identification with S-transform—GST—and Convolutional Neural Network. They apply a GST on the ECG signal and generate an image from each row. Each row generates a trajectory where the real part is on the x-axis and the imaginary part is in the y-axis. All the trajectories from each row generate an overlapped image that feeds the Convolutional Neural Network. They designed their own CNN model and they use three databases for training, testing and validation of their model. The best result on identification has an accuracy of 96.63% and do not report experiments on authentication.

2.7 Multibiometrics

Biometric systems can use a combination of sensors, algorithms, samples, instances (e.g. left and right iris) or modes (i.e., biometric traits such as fingerprints and ECG). Such systems are said to implement multibiometric methods [24].

Multibiometrics has some advantages: it enhances a unibiometric (single biometric trait) system, reduces the probability of the system encountering two users with the same biometric information (non-universality characteristic) and can speed up indexing for identification in a large-scale database. The use of multi trait can narrow down the amount of potential matching candidates and focus the search only among a few stored templates. In addition, multibiometrics provides robustness under noisy environments; if one trait fails, we can use another trait without disrupting the biometric process [24].

A multibiometric system follows a similar flow as a unibiometric system (see Figure 1), except that fusion occurs at the matcher or the decision modules.

2.7.1 Multibiometric Classification

A multibiometric is classified according to the source of information that will be used in the fusion [24]. A multibiometric system acquires the information to fuse from multiple sensors, algorithms instances, samples, modes and can mix between these in a hybrid approach. Based on these sources of information, a multibiometric system can be Multi-sensor, Multi-Algorithm, Multi-Instance, Multi-Sample and Multi-Modal or Hybrid.

Multi-sensor biometric is when two or more sensors capture the same biometric trait. As an example, a multi-sensor biometric system can have a 2D camera and a 3D camera. They capture the same trait (face) but they provide more characteristics [75].

A system that has different algorithms that evaluate the same trait is a multibiometric system known as multi-algorithm biometric. As an example, a multi-algorithm biometric can be an ECG authentication system that uses SVM and KNN for classification [76].

Multi-instance biometric has two or more data of the same trait class. For example, we can implement multi-instance biometric by using left index finger and right index finger. Both of them are fingers, but they are from different extremity [77].

Multi-samples biometrics uses the same biometric sensor to take two or more samples of the same biometric trait. For example, a camera that takes the left, front and right side of the face [24].

Multi-modal biometrics uses different traits in order to identify a person. For instance, a multi-modal biometric combines Fingerprint + ECG + Face + Iris + Voice in order to perform identification or authentication. It is called Bi-modal if only two biometric traits are used [41].

A hybrid biometric combines any of the previous sources of information in order to provide one hybrid multibiometric solution [78]. As an example of a Hybrid Biometric

solution is a multimodal biometric system that uses ECG and Fingerprint with an ECG multi-algorithm classifier that implements SVM and Neural Networks.

2.7.2 Multibiometric Fusion

Fusion in a multibiometric process occurs before or after matching [24]. Fusion that take place before the matching stage can be divided into two types: sensor level fusion (multi-sensors) and feature level fusion (multi-samples) [24]. Fusion that takes place after the matching stage has three categories: match score level fusion, rank level fusion and decision level fusion. These last three categories of fusion are part of the fusion of classifiers. One more category is the dynamic classifiers; they fuse biometric traits to narrows down the results in order to improve the speed of identification. Figure 5 shows these fusion levels

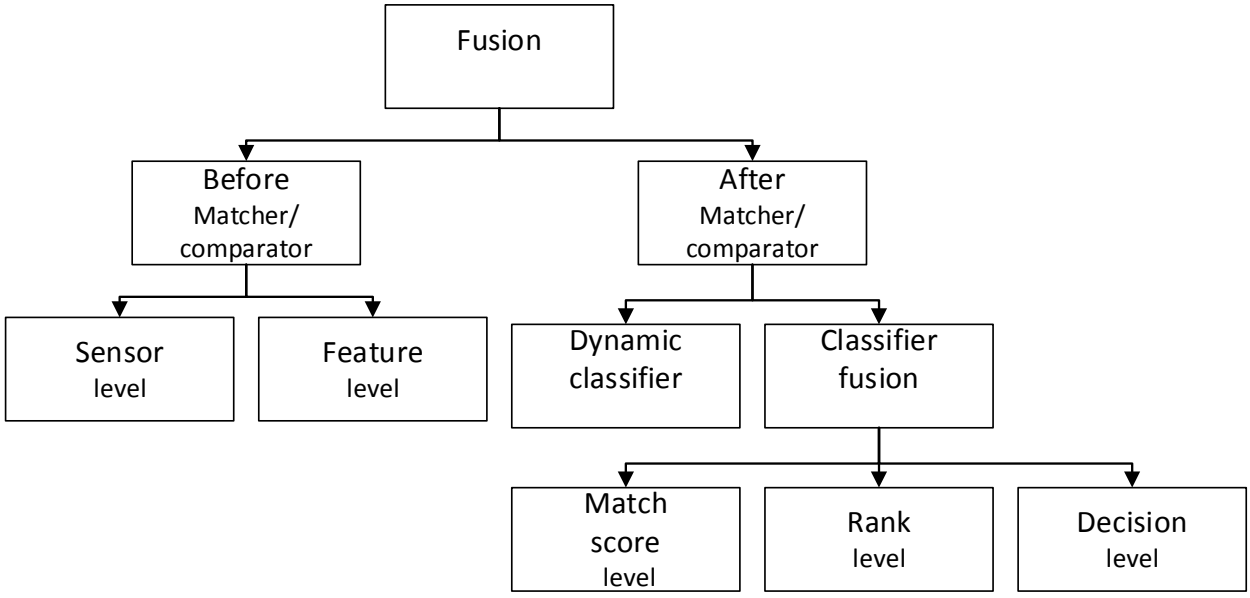

Figure 5. Fusion Levels of a Multibiometric System adapted from [24]

Sensor level fusion combines several samples of the same biometric trait from different perspectives or take samples of the same biometrics trait with different sensors. As an example, a multi-sensor biometric system can have a 2D camera and a 3D camera. They capture the same trait (face) but they provide more characteristics [24].

Feature-level fusion normalizes, transforms and reduces features from multiple biometric algorithms. Feature-level fusion fits all the features into the same vector space to feed a matcher and yield the result [27].

Match score level fusion combines the match scores of different biometric to generate a unique match score.

Rank Level fusion narrows down the identification possibilities with one biometric trait and confirms the identification with the second biometric trait [79].

Decision Level Fusion uses unibiometrics results (Match or non-match) to indicate if a final match has been found. Fusion at the decision level does not need the features or scores of the classifiers [80]. This is an advantage when we are combining proprietary algorithms into the biometric system. In such scenario, we solely have access to the decision as opposed to the score [24]. Some decision level fusion approaches are Majority Vote Fusion, Linear Weighted Fusion and Behavior Knowledge Space Fusion.

Majority vote is the most used one; it returns the decision that the majority of classifiers choose [81].

Linear weighted fusion is an extension of the majority vote. It assigns a weight to the matchers based on their performances [81]. Atrey *et al*. [82] indicates that weight assignment is the major drawback of linear weighted fusion.

Behavior Knowledge Space is a statistical fusing rule that calculates the probability of a decision (match or non-match) based on the combination of results from the individual matchers. The decision with higher probability in a specific combination of results is the final decision for the multimodal biometric [81]. This technique performs effectively for multi-class (decisions and matchers) problems where large datasets are available for training [83].

2.7.3 Multimodal approaches using ECG

While research on bimodal or multimodal biometrics is vast; currently, limited research exist on the combination of ECG and Fingerprint biometrics. This is possible because ECG is a relative new topic in biometrics. The found literature are described in the following paragraphs.

Manjunathswamy et al. [84] used ECG and fingerprint in a bimodal biometric system. They combined the features from ECG and fingerprint and authenticated or identified a subject when 11 out of the 14 features were above a threshold of 75%. They obtained an FRR of 0% with a FAR of 2.5%. However, they collected ECG and fingerprint data separately and provided no information on the number of users in the study. They perform fusion at the feature level. They extract 12 ECG features and 2 fingerprint features. They set a threshold of 75% on all these features. The algorithm reaches a match when 11 out of the 14 features pass the threshold.

Sing et al. [15] presented a study that shows the feasibility of using ECG as a biometric. In the same study, they evaluated a bimodal biometric system that used ECG and fingerprint traits on 73 subjects. They collect ECG and fingerprint data separately: the ECG records are obtained from the Physionet database [85] and the fingerprints were collected from NIST-BSSR1 [86]. The NIST-BSSR1 database does not provide fingerprint data; instead, it provides matching scores from the NIST Bozorth3 matcher. To perform fusion, they applied a score transformation using a weighted sum rule, and then, they used a threshold for matching. They reported an EER of 1.52%.

2.8 Continuous Wavelet Transform

There are many domain transformations to analyze signals. Fourier transform is one of them and changes the signal from the time domain to the frequency domain. Many applications require frequency analysis and use this transformation—e.g. signal filters. The main problem that Fourier transform has, is the missing time component [87]. With a Fourier transform, we are able to know what frequencies are present in the signal but we do not know

the time location of those frequencies. Short Time Fourier Transform—STFT—splits the signal in smaller time windows to add the missing time component and then applies the Fourier transform in each window [87]. The size of the window is proportional to the time and frequency resolutions. A smaller window provides better time resolutions but poor frequency resolutions. A wider window provides better frequency resolutions but poor time resolutions [88]. This trade-off is the result of a physical phenomenon known as the Heisenberg uncertainty principle [89].

Multi-resolution analysis—MRA—adjust the frequency resolution according to the frequency. In contrast to STFT, MRA does equally resolve each frequency component. MRA has good time resolutions with poor frequency resolutions at higher frequencies and low time resolutions with good frequency resolutions at lower frequencies[89]. MRA does not solve the Heisenberg uncertainty principle but the frequency resolution flexibility provides a better solution than STFT. The frequency flexibility is an advantage because—in general—signals have more low frequency components than high frequency components. Higher frequency components are usually present for just a short period, while lower frequency components are present most of the time [90].

Wavelet Transform is a multi-resolution analysis. The term wavelet means small waves and that is what they do; they use many small waves to transform the signal [91]. There are two types of wavelet transformations known Continuous Wavelet Transform—CWT—and Discrete Wavelet Transform—DWT. CWT can be processed on digital computers but is computational more expensive than DWT. CWT provides a better solution for applications that require a more detailed analysis of the signal in the time and frequency domain. On the other hand, DWT provides a better solution for signal compression, denoising or transmission [91].

CWT transformation is similar to STFT; both of them multiply the signal by a function but CWT uses a wavelet function $\psi(t)$ instead of sin and cos. In CWT, an input signal $x(t)$ is multiplied with the wavelet function ψ at the location $\tau = 0$—beginning of the signal. This result is integrated over all times of $x(t)$ and then multiplied by $\frac{1}{|s|^{\frac{1}{2}}}$. This last multiplication

normalizes the energy of the signal. For the next step, we displace the wavelet function ψ a τ value and the operation repeats until τ reaches the end of $x(t)$. This operation is a convolution of the wavelet ψ and gives the time component of the CWT. We do the first convolution with the mother wavelet. A mother wavelet is a wavelet that is not expanded or contracted, in another words it has a scale $s = 1$. The next step, repeats the same procedure as before but with a different scale value. This operation with different scale values provides the frequency related components of the CWT. In CWT, we do not refer to frequencies because we use scales and a scale value used is $s = {}^{1}/_{freq}$. All the CWT transform is represented in (3).

$$\Psi_x^{\psi}(\tau, s) = \frac{1}{|s|^{\frac{1}{2}}} \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t - \tau}{s} \right) dt \tag{3}$$

We can use many wavelets in the CWT; one of them is the Morse wavelet. This wavelet works well with time localized applications [92]. The Morse Wavelet is represented in (4).

$$\Psi_{\beta,\gamma}(\omega) = \int_{-\infty}^{\infty} \Psi_{\beta,\gamma}(t) e^{-i\omega t}\, dt = U(\omega) a_{\beta,\gamma} \omega^{\beta} e^{-\omega^{\gamma}} \tag{4}$$

Where, $U(\omega)$ is a step unit function, $a_{\beta,\gamma}$ is the normalization constant and γ and β controls the shape distortion of the wavelet. γ alters the symmetry and β the time decay. Other well-known wavelets are Morlet [93], Mexican Hat [94], Poisson [95]. The usage depends on the requirements of the application.

DWT follows a similar procedure as CWT, but the main difference is the scaling parameter [91]. Scaling in CWT is based v octaves $2^{\frac{j}{v}}$ and DWT use a scale with base two 2^j, where $j = 1,2,3, \ldots$. This scaling affects the computational cost of CWT because more scales requires more processing.

2.9 Deep Learning and Convolutional Neural Networks

Deep learning is a machine learning technique that automatically determines and extracts features in order to perform classification [96]. They do not require feature engineering like traditional machine learning because multiple layers interpret data at different abstraction levels. This data interpretation is the automatic feature engineering that will determine the features for classification [96].

Convolutional Neural Networks or better known as CNN is one of many successful approaches in Deep learning. CNN is a powerful version of a feed-forward neural network and the main application is image recognition and classification [21]. Although, images is the main applications of a CNN, it is possible to use CNN in different classification applications such as diagnose with physiological signals, text processing among others [97]. CNN require large datasets in order to perform properly. This was a limitation until the appearance of large datasets like Open Images dataset [98]. The appearance and fast development of GPU processing is another factor that has influence the increase popularity of CNN [99].

There are many variants of the CNN architectures; however, all of them use three types of layers known as Convolutional, pooling and fully-connected layer [21]. The convolutional layer performs convolutions on the image pixels in order to detect features. The convolution kernel is a small matrix of weights that convolves—slide—the image pixels; as a result, each convolution generates one pixel. The collection of all the generated features forms a feature map. Usually, models have more than one kernel and generate several feature maps. These generated features maps will create another feature maps in a deeper layer. The number of layers depends on the architecture design of the model. Convolution is a linear operation but many applications are non-linear. An activation function is after each convolution operation to solve the non-linear problem. One of the most used activation functions is ReLU—Rectified Linear Unit—and replaces all negative values of the feature map with a zero [100]. Other activations functions that solve the non-linearity problem are $tanh$ and $sigmoid$ [101] [102].

The goal of the pooling layer is to prevent errors when images represent the same object in different positions or angles—shift-invariance. To accomplish shift-invariance, the pooling layer reduces the dimension of the features map—known as spatial pooling—while retaining the most important information. A pooling layer is between two convolutional layers and they obtain the spatial pooling from the features map of the previous convolution layer. They use a window among the features map and gets the most important value. They are several approaches to determine the most important value; one is the maximum value of the window—max pooling [103]—or the average of the windows values—Average pooling [104]—or just the sum of all values of the window—sum pooling [105]; where most applications use max pooling [21].

The fully connected layer goal is to learn from features of previous layer. All the neurons—features—of the previous layers connect to neurons in the next layer—fully connected. These features create a non-linear combination that provides more information for classification. This layer is not always necessary, but it can provide high-level features for a better classification [106].

The last layer is the classification layer or output layer; we did not mentioned earlier because a CNN can be a feature extractor only [107]. In this case, the classification layer does not exist. When the purpose of CNN is classification, this layer has an activation function—usually a softmax operator [108]—that calculates the probability that the input has in each class. The sum of all class probabilities is always one.

2.10 Signal Processing Tool and Evaluation Method

As part of this thesis, we have developed a signal-processing tool that detects the R peak of an ECG signal. An R peak detector with a low average time error is important to reduce the distortion on time location of the R peak. If the time distortion is low, it can help the authentication algorithm to have better templates that improves the accuracy.

Some multibiometric approaches use more than one threshold to evaluate results. However, the literature presents evaluation procedures with one threshold only. Therefore, in this thesis we formalize a method to evaluate multibiometric with multiple thresholds.

Section 2.10.1 and section 2.10.2 present a background on the tool and method that we develop as part of this thesis.

2.10.1 ECG R-Peak Detector

ECG biometrics requires the detection of fiducial points. These fiducial points are distances that we measure from the R peak to the P, Q, S and T peaks. R peak is the reference for all the fiducial points; therefore, accuracy is important in the detection of R peaks. In order to use QRS detectors in many applications with different scenarios, they should be accurate, fast and capable to detect fiducial points within noisy signals [11]. A previous work by Arzeno *et al*. [109] analyzed five QRS detectors. The method that achieves the best results in terms of Average Time Error is the method based on Hilbert Transform with the second derivative. However the sensitivity and positive predictivity are inferior compared to other methods such as Hilbert transform with automatic threshold, Hilbert transform with secondary threshold, Hilbert transform with squaring function (correction of polarity) with patient-specific threshold and Hilbert transform with squaring function with automatic threshold.

Another approach in QRS detection was presented by Benitez et al.[110]. They used the Hilbert transform as a filter of the first derivative signal. They detect the peaks with the help of a variable threshold; they modulate the threshold for each window of data in order to perform more localised test on the values of the signal. They use a window size of 1024 samples and calculate: the Root Mean Square (RMS) and the Maximum Value (MV) for each window. If the RMS value of the current window is larger or equal than 18% of MV of the current window, then the threshold value is set to 39% of MV of the current window. If MV in the current window is larger than twice MV of the previous window, then the threshold is 39% of MV of the previous window. If the RMS value in the current window

is less than 18% of MV of the current window, then the threshold is set to 1.6 times of the RMS value of the current window.

Arzeno et al.[109] combines and evaluates some of the previously described methods. The first method they tested (Method I) is the Hilbert transform with the variable threshold introduced by Benitez et al.[110]. The second method (Method II) is the Hilbert transform of Benitez et al. [110] with a secondary threshold that was presented by Hamilton and Tompkins [111]. The third method (Method III) tested was the Hamilton Tompkins [111] and the fourth (Method IV) is the Hamilton Tompkins [111] technique with the variable threshold of Benitez et al. [110]. Finally, the fifth method (Method V) evaluated was based on the work of Benitez et al. [110], but used the second derivative instead of the first. The results of this study showed that Method III results have the highest sensitivity and positive predictivity; however, it also produces the highest average time error. Method V was found to produce the lowest Average time error; however, Method V achieves the poorest results in terms of sensitivity and positive predictivity.

All of these methods comply with the common structure of QRS detectors. A common QRS detector has a preprocessing stage, composed by Linear Filtering and Non Linear Filtering; and a Decision Stage, composed by peak detection Logic and a Decision process [112].

When ECG biometrics uses time-based features, it requires a low average time error. QRS detectors techniques and filters displaces ECG signal in the time domain. This displacement in time introduces an error in the detected features that are measured by the average time error. A QRS detector with a low average time error provides a more accurate measure of the ECG features; this has a direct effect in the accuracy of the ECG biometric. Therefore, in order to achieve better results with ECG authentication, part of this work develops an R peak detection tool with the lowest Average Time Error. Section 6.1 describes this R peak detection tool in more detail and section 6.2 presents the evaluation of this tool.

2.10.2 Calculation of DET and EER with Multiple Thresholds

As mentioned in section 2.3.2, the DET curve is used to calculate the EER in a unibiometrics system; however, some multibiometric systems uses it as well [113]. In a multibiometric system He *et al*. [114] fuses at the score level in order to use the same evaluation method as in a unibiometric system. Fusing at the score level reduces the different scores of a multibiometric system into a one score. With one score, the adjustment of a single threshold parameter is enough to obtain the performance of the system in the DET curve. Chen *et al*. [115] presents a multimodal biometric that uses two biometric treats: Face and IRIS. They perform the fusion at the match score level which reduces to a one score [115]. The authors extract biometric features from the Face and Iris. Then, the team classifies these features with a Wavelet Probabilistic Neural Network (WPNN). The algorithm reaches a decision based on Probabilistic average that goes beyond or under one specific threshold. Brunelli *et al*. fuses face and voice at the score level with five different operating matchers and reduce to a single score [116]. Hong *et al*. fuse at the measurement level and use face and fingerprint biometric traits [117]. Dieckman *et al*. fuse at the score level biometric traits like voice, lips motion and face [118]. Another study [119] uses fingerprint, iris and face to have a multibiometric solution and fuses at the feature level. Kim et al. propose fusion at the score level of face and speech biometric traits, the goal is to reduce complexity in order to apply it online [120].

In unibiometrics system, the DET graphs are direct because only one threshold is involved. In multibiometrics, many studies present works at the match score level fusion. When fusing at the match score level, it will generate one general match score. With one match score, one threshold is enough to plot the DET curve. Fusion at the decision level generates a problem when plotting the DET curve. At the decision level fusion several thresholds values are available, but DET is a graph that uses only one threshold value. In order to solve this problem, section 7.1 of this work presents an approach to estimate the DET graph in a multibiometric system with several thresholds. Section 7.2 presents the evaluation of this proposed approach.

Once the DET graph is obtained, we can calculate the EER. But, to the best of our knowledge, the only existent approach for calculating the EER is visual or using the approach by Poh et al. [121]. They calculate the Equal error rate as:

$$EER = \frac{1}{2} - \frac{1}{2}\text{erf}\left(\frac{\text{F-ratio}}{\sqrt{2}}\right) \quad (5)$$

Where,

$$erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$$

And,

$$\text{F-ratio} = \frac{\mu_{genuine} - \mu_{impostor}}{\sigma_{genuine} + \sigma_{impostor}}$$

Where, $\mu_{genuine}$ and $\mu_{impostor}$ are the mean for the genuine and impostor scores. $\sigma_{genuine}$ and $\sigma_{impostor}$ are the standard deviation for the genuine and impostor scores

This last approach only applies to normal distributions. In biometrics, distributions are not normal [122], therefore we need an alternative method to calculate mathematically the EER. As part of the estimation of DET, section 7.1 of this work also presents an approach to calculate the EER for distributions that are not normal. Section 7.2.2 present the corresponding evaluation of this approach.

2.11 Conclusion

In this chapter, we have introduced many concepts that we use in this thesis. We started with an overview of biometrics and explained the most used biometric traits. We add more details for fingerprint and ECG biometrics because we will be using them in this Thesis. We present information related to Continuous Wavelets Transforms, SVM, CNN. We also

presented the background for a QRS detector that we are using in this theses and an evaluation method for multiple thresholds.

Previous studies have shown that the ECG trait can be used as biometric [13], [29]–[31]. ECG biometric methods do not reach the performance level of fingerprint approaches if the ECG signal length is kept short. Nevertheless, ECG features are more difficult to spoof.

In contrast with existing work, in this thesis we will develop ECG authentication algorithms with short acquisition time and high accuracies. Related works in this section confirm that an automatic threshold with SVM provides better results. In addition, the use of deep learning can help to extract better features to improve the accuracy. The section of multibiometric, confirm that these systems can perform better than unibiometric ones. Although the literature is quite limited on bimodal biometric approaches with ECG, we will provide an algorithm that use ECG and fingerprints to improve the accuracies of the related works.

In order to evaluate the bimodal system with more than one threshold, we will use our method to estimate the DET curve with the mathematical approach to calculate the EER. In addition, the algorithms with ECG will use our QRS detector of an ECG signal that is fast, accurate and it can process noisy signals with low average time error.

Chapter 3.

ECG Authentication with SVM

The ECG process employed in this work is adapted from our previous work [11]. However, we have replaced the matcher module to account for the bimodal nature of the current system. Therefore, we propose a matcher that uses an SVM classifier with an RBF radial basis function kernel.

3.1 Design of ECG with SVM

Our previous work [11] presented a unibiometric approach that requires associating each feature with a threshold. Therefore, such approach entails a tedious search for the optimal combination of all thresholds. A multi-biometric system further exacerbates this issue by increasing the number of thresholds. Hence, although using the matcher from [11] is possible, SVM based matching for the ECG features is advantageous in a multibiometric context because it eliminates the thresholds associated with the ECG biometric method and eliminates the tedious task of threshold tuning. ECG authentication features are parametric where SVM performs better and does not suffer from over fitting. After a series of experiments, we determine that RBF SVM kernel (Radial Basis Function) performs better for ECG

authentication features. Furthermore, section 3.2 evaluates the matcher from our previous work with the one using SVM.

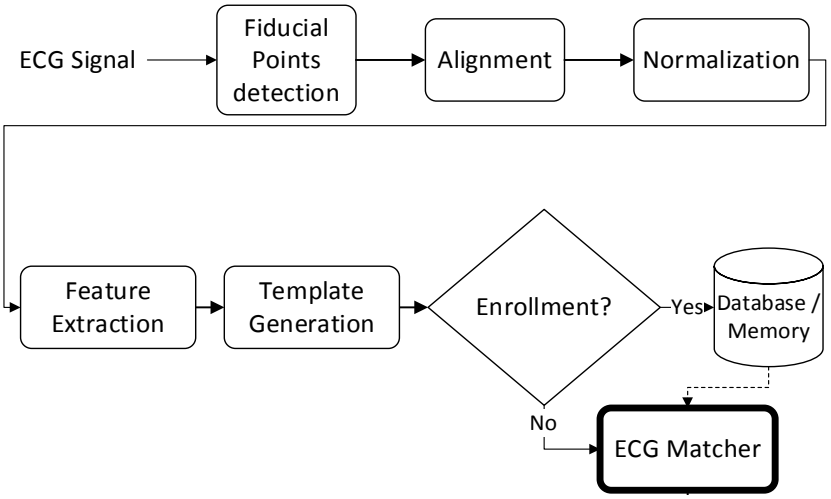

Figure 6. SVM ECG Authentication Diagram.

In accordance with our previous work [11], we use 4-second long signals for the ECG authentication. Such a signal length typically includes several complete heartbeats. The number of heartbeats depends on the heart rate of the subject during signal collection. Figure 6 shows our ECG biometric method which consisted of the following steps:

1. Detect the fiducial points (i.e., QRS complex, LP & TP valleys and P & T peaks) from each heartbeat in the ECG signal;

2. Extract each complete heartbeat from the signal and align it around a reference point, (in this case the R peak) (Figure 7).

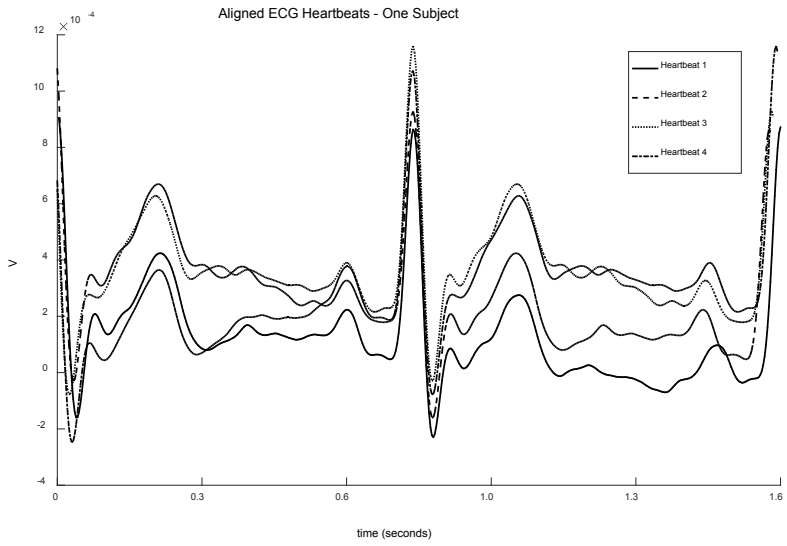

Figure 7. ECG Heartbeat Alignment [11].

3. Normalize each ECG heartbeat to eliminate disparities caused by heart rate variations by dividing each feature by the total length of a heartbeat [10];

4. Extract the ECG features as previously described [11] (Figure 8) to obtain eight features that represent the distance in time between peaks and valleys and amplitudes distances;

5. Generate the template, a set of all features extracted in step 4, necessary for enrollment or authentication;

6. If the system is enrolling a user, the template from step 5 is stored in a database or memory. If the system is authenticating a user, the template is fed to the ECG matcher that then compares it with the stored template. The SVM classifier running within the matcher produces a binary result of match or non-match. SVM sets a boundary around the *N*-dimensional training set data points collected during enrollment (where *N* corresponds to the number of features). The algorithm considers any input template that corresponds to a data point (*N*-dimensional space) within the boundary to be a match; otherwise, it is a non-match.

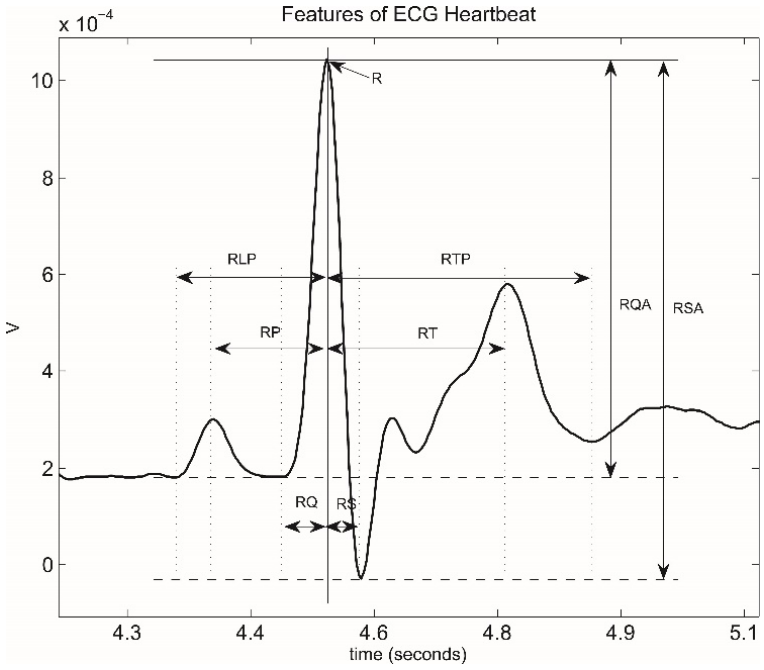

Figure 8. Features Extracted From ECG [11].

Because this work focuses on authentication, we want to determine if certain biometric features correspond to a genuine user (one on one match). Therefore, we train the SVM classifier on a single class since we only collect data for genuine users during enrollment; we cannot obtain data for all the possible impostors. Hence, we solve the classification problem with a one-class SVM classifier.

The arrangements of features in the SVM space makes it difficult to accurately classify them with a linear kernel or polynomial kernel. In our preliminary tests, we considered using the linear kernel, polynomial kernel and Radial Basis Function (RBF / Gaussian) kernel. The RBF kernel performed better for our ECG authentication approach. We are using time and amplitude ECG features that result in an SVM space where the RBF kernel produces a boundary that does not seem to under or over fit the training set.

The RBF kernel that we use is:

$$K(x, xi) = e^{-\|x - x_i\|^2} \tag{6}$$

where, x is the feature vector stored in the template and x_i is the feature vector corresponding to a user attempting authentication. We place the kernel (equation (6)) in the general SVM classifier equation to obtain equation (7):

$$f(x) = \text{sgn}\left(\sum_{i=1}^{n} \alpha_i e^{-\|x-x_i\|^2} - \rho\right) \tag{7}$$

where α_i is the Lagrange multipliers (calculated when solving the minimization formulation of SVM); ρ is a soft margin parameter that controls how wide is the classification margin.

The SVM classifier in equation (7) provides a matching result based on the features extracted at the input and the features stored in the database. The system sends the result to the bimodal authentication algorithm to perform fusion at the decision level.

3.2 ECG with SVM Evaluation

We compared this proposed ECG-SVM matcher with our previous ECG-Threshold matcher using 4 Physionet databases [123]: MIT-BIH Arrhythmia Database [124], MIT-BIH Normal Sinus Rhythm Database [123], European ST-T Database [125] and QT Database [126]. For this evaluation, we randomly selected a total of 73 ECG records from the databases. We divided the signal of each subject into seven fragments of data: the first fragment was 60 seconds in length and was used for enrollment and the six remaining fragments were 4 seconds in length and were used for authentication. Each fragment corresponded to a random portion of the ECG record. We did not set a threshold for evaluation because ECG-SVM yields a matching result (genuine or impostor) and not a score; therefore, we only had one operating point and do not generate a DET graph. We measure the number of true positives (TP) and false positives (FP) to calculate the TAR and FAR.

We evaluated the best performing algorithm by measuring the TAR when both algorithms had the same FAR value. The ECG-Threshold matcher [11] reported a TAR of 84.93% with a FAR of 1.29%. The ECG-SVM matcher only has one operating point, therefore the TAR and

FAR cannot be calibrated. The ECG – Threshold matcher has several operating points (see Figure 9), therefore we can adjust the thresholds of this algorithm to generate a DET graph. The DET graph for the ECG-Threshold matcher shows that a FAR of 7% has a FRR of 4.5%. This FRR represents a TAR of 95.5% ($TAR = 100 - FRR$). The ECG-SVM achieves a TAR of 100% with a FAR of 7%. Therefore, for a FAR of 7%, the ECG-SVM performs better than ECG-Threshold.

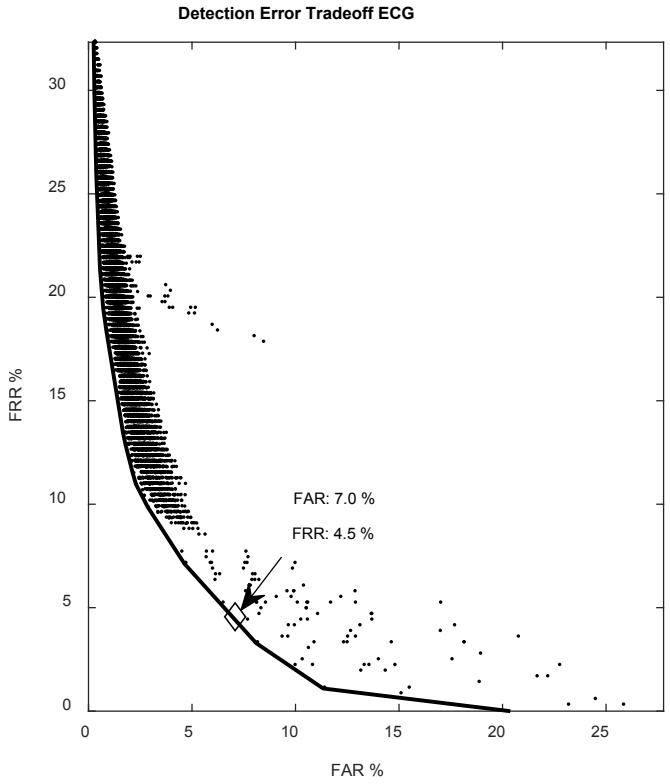

Figure 9. DET Approximation for ECG-Threshold Algorithm

Chapter 4.

ECG Authentication with Deep Learning

In this chapter, we will explain the development and evaluation of an ECG authentication algorithm that uses a hybrid approach. It uses a pre-trained deep learning model to extract features automatically from an ECG signal. Later, takes these features and perform authentication with an SVM one-class classifier.

4.1 Design of ECG Authentication with Deep Learning

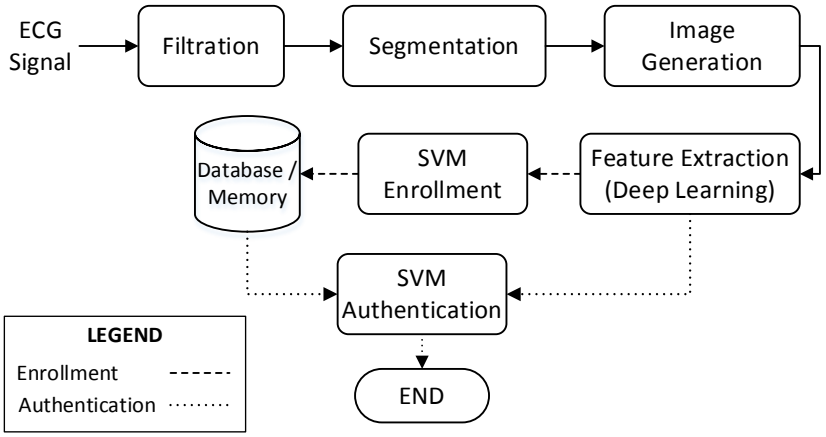

Figure 10. ECG authentication algorithm with deep learning.

The general view of the ECG authentication algorithm with deep learning follows the diagram presented in Figure 10. Where the input is an ECG signal. The filtration removes the noise of the signal ECG signal and prepares it for the next stage. After the signal is clean, it goes to a Segmentation stage where it detects fiducial points to separate the ECG signal by heartbeats. The image generation stage generates a heartbeat signal image that contains information of frequency, time and gain. The feature extraction stage uses a pre-trained Convolutional Neural Network (CNN) model to get unique features from the generated images. If the user is in enrollment mode, these features train an SVM model and store the model in memory/database. If the user is in Authentication mode, the SVM uses these features to perform a classification (genuine or impostor) based on the model stored in memory/database. The following subsections provide more details of each stage.

4.1.1 Filtration

The first stage is filtration, where it removes baseline wander and power-muscle noise of the ECG Signal. To remove the baseline wander, this work uses a Polynomial Fitting method instead of Linear Filtering. Linear Filtering produces a signal distortion that affects the original ECG signal. A filter with a higher accuracy in baseline wander removal has a higher distortion of the signal [127]. Polynomial fitting might not be as accurate as linear filtering, but it does not distort the original signal. This characteristic is important for an accurate ECG authentication and we should avoid linear filtering to the utmost. This work uses the Least Square Fitting method to find the best coefficients of the polynomial. A

heuristics experiment in this work finds that a 6th degree polynomial achieves the best results. The general equation of the polynomial used in this work is presented in (8), where seven coefficients represent a 6th degree polynomial.

$$p(t) = p_1 t^6 + p_2 t^5 + p_3 t^4 + p_4 t^3 + p_5 t^2 + p_6 t^1 + p_7 t \quad (8)$$

A subtraction of the polynomial (8) from the contaminated ECG signal represents an ECG signal without baseline wander noise. See (9).

$$ECG(t) = ECG_c(t) - p(t) \quad (9)$$

Where, $ECG_c(t)$ is the ECG signal contaminated with baseline wander noise and $ECG(t)$ is the clean ECG signal after the removal of the baseline wander noise.

After baseline wander removal, the next filter removes the noise from the muscles and power lines. To remove these noises this works uses a Savitzky-Golay Smoothing Filter. As previously mentioned, it is important to avoid linear filtering because of the distortion of the signal. The Savitzky-Golay filter is a polynomial fitting filter, which prevents the distortion of the original signal [128]. The characteristics of the Savitzky-Golay filter are of order 0 and a frame length of 9. This configuration filters the signal four times in order to provide a better filtration while maintaining the unique characteristics of the ECG signal.

4.1.2 Segmentation

An ECG signal has fiducial points that determine the biometric features of an ECG signal. Traditionally, these features generate a template that will be match when authentication is in progress [11]. However, this work uses fiducial points as a reference to segment the signal by heartbeats. Two fiducial points delimits the beginning and the end of a heart beat: LP and TP. See Figure 11.

Detection of the R peak is necessary to detect LP and TP fiducial points. An algorithm based on differentiation [129] detects the R peak. This algorithm is preferred due to its fast performance, does not require a threshold calibration and has a low average time error. With the R peaks as a reference, we use the algorithm presented in [112] to detect the LP and TP fiducial points.

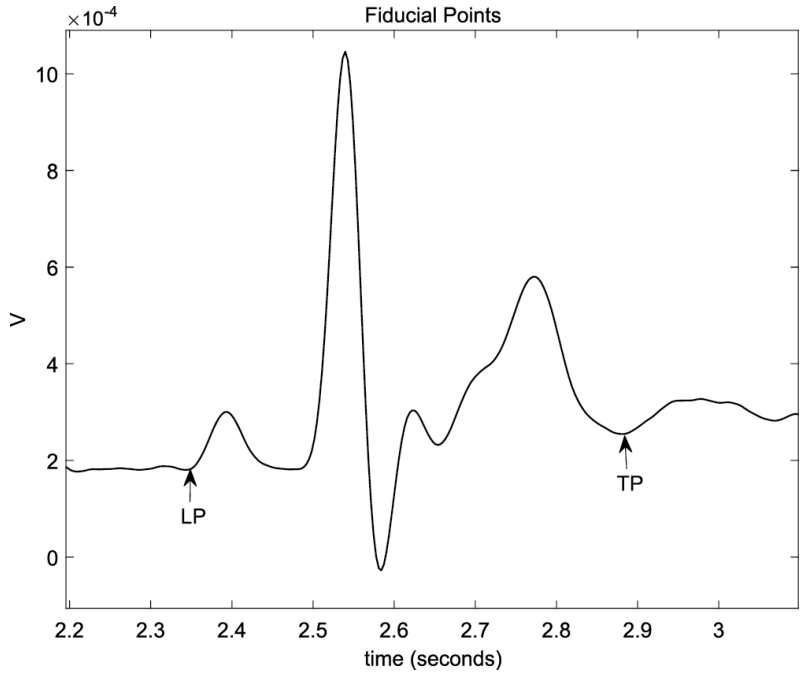

Figure 11. ECG heartbeat segment.

LP and TP fiducial points segment each heartbeat across the ECG signal. These heartbeats are later stored in vector (10) that will generate the images in the next stage.

$$ECG(t)_S = \left[ECG(t)_{HB_1}, ECG(t)_{HB_2}, \cdots, ECG(t)_{HB_{n-1}}, ECG(t)_{HB_n}\right] \quad (10)$$

Where, $ECG(t)_S$ is the segmented array that contains n number of ECG heartbeats.

4.1.3 Image Generation

This work generates an image from ECG in order to feed a CNN model in the next stage. A traditional ECG image presents features based on time only—See Figure 11.

However, literature shows that other biometric features exist in the frequency domain [130]. In order to present an enhanced image with enhanced information, this work merges time and frequency components through a wavelet transformation to generate an image.

The transformation process uses the segmented ECG signal and converts each heartbeat in a JPG image. This image has time components across the horizontal axis and frequency components—better known as scale in wavelets—across the vertical axis. The colors in the image represent the normalized gain component.

This work uses Morse Wavelets because it works better with time localized applications [92]. Previous works on ECG authentication has shown that time based features are clearly located—based on fiducial points—in the signal [11], [41]; therefore, it is important to use a wavelet that delivers better time resolutions over frequency resolutions. Frequency features of an ECG signal are also important but they are non-fiducial [130]; therefore we use a wavelet that is more permissive with frequency resolutions. We use a value of gamma $\gamma = 3$, where γ represents symmetry in time. Positive or negative skew affects the resolution in time; therefore; no symmetry means no time resolution affection. That is why 3 represents the best resolution in the time domain and is used in most applications [131]. The decay parameter that we use is $\beta = 20$, because this value represents the low decay in time and frequency, which leads to better frequency resolution without affecting the time resolution [131]. Another parameter is the number of voices per octave and is set to 12. This parameter controls the scale—number of frequencies to analyze—of the wavelet. Twelve is the limit value to analyze the ECG signal. After analyzing different images from ECG signals, we conclude that more than twelve does not provide more information; the information is either redundant or has no information—Zero gain.

The previously described settings are the parameters to perform a wavelet transform to the segmented ECG signal. Each segment represents a heartbeat and we generate an image for every heartbeat—see Figure 12. The wavelet transformation provides coefficients—they have a real and imaginary part—and the modulus of these coefficients provides the gain of a specific frequency at a specific time. A range of colors represents these gains in

the image. The chosen range of colors is from RGB (62, 38, 168) to RGB (249, 251, 20) and we divided in 256 colors between the range—see the color bar in Figure 12. This specification is recommended because avoids misinterpretation of high and low values when they are represented by dark colors [132]. The x-axis of the image represents the time and the y-axis is the frequency. The y-axis has a logarithmic scale because most of the information for ECG is located in low frequencies. A logarithmic scale shows more detail information for lower values—frequencies in this case. Please note that this work generates the images based on time and frequency in the x-axis and y-axis, respectively. We generated images like that for a better understanding of the graphs. Normally, wavelets display information in terms of time and scales and not in time and frequency. Scales represents the range of frequencies that each wavelet covers, a higher scale value represents a lower frequency and a lower scale value represents a higher frequency value.

All the generated images have a fixed size of 224pixels by 224 pixels—see Figure 12—because the pre-trained CNN model for feature extraction requires this size of images—next section will present more details about it. Fixed-size images are not only for feeding the CNN model; they implicitly normalize the ECG signals that they represent. Normalization is necessary because physical activities or change of emotions produces changes in the heart rate that affects the performance of ECG authentication [10]. A fixed-size image automatically adjusts the length of the heartbeat to the same size. A heartbeat might last 1.3, 1.5 or 1.8 seconds; any duration has to fit in 224 pixels. The x-axis values 0 to 1.3 seconds, 0 to 1.5 seconds or 0 to 1.8 seconds will always be a value of 0 to 223 pixels. It is important to mention that the generated image does not have labels of x-axis or y-axis; it is just the ECG image with gain —normalized modulus, frequency and time components.

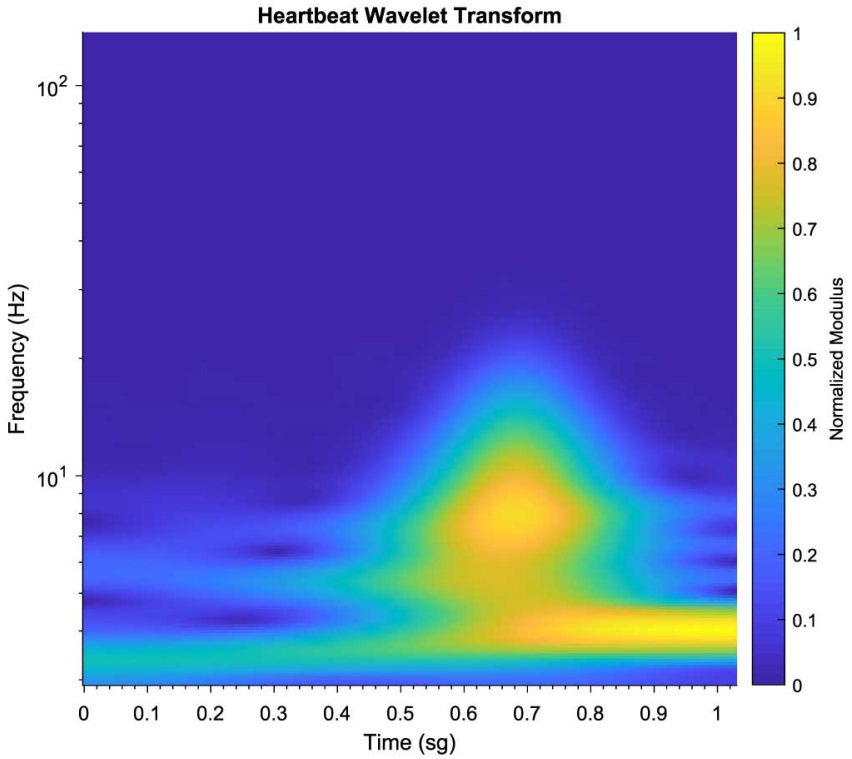

Figure 12. Generated Image from Wavelet Transform of one Heartbeat. The algorithm uses the image with no labels. Labels in this figure help to interpret the image information.

All the generated images are stored in a vector of images (11), where each element corresponds to a heartbeat. This vector is similar to the segmented vector (10); the difference is that vector (11) is in a different domain, it changed from the time domain t to the image domain img. The subscript s indicates that is part of an ECG segmented by heartbeats.

$$ECG(img)_S = \left[ECG(img)_{HB_1}, ECG(img)_{HB_2}, \cdots, ECG(img)_{HB_{n-1}}, ECG(img)_{HB_n}\right] \quad (11)$$

4.1.4 Feature Extraction with Deep Learning

Millions of images have trained several deep learning models and CNN has highly accurate results for image classification [133]. One of the key aspects of CNN deep learning is that they have strong automatic feature detection algorithms. These automatically detected features are fundamental part for the image classification [133].

This work uses GoogLeNet [23] as a deep learning model to extract features from the ECG generated images. Enrollment or authentication stages will use these extracted features to train or validate users with SVM; section 4.1.5 describes it with more details.

We use pre-trained deep learning models because they use millions of images to generate them and have been well benchmarked against each other [133]. One challenge that pre-trained models face is that researchers designed them for image classification but ECG signals are not images. That is why some of the related works have trained and create their own deep learning models for ECG authentication; however, their training data is not comparable with the training data used in pre-trained models. Our solution to this challenge is to generate images from ECG signals as previously described in section 2.8.

Another challenge is that deep learning is for classification tasks. In biometrics, authentication is a special classification case called one-class classification—also known as anomaly detection— that only trains the model with data from genuine users. Any other data that does not fit the model is an impostor. Deep Learning does not handle the one-class classification case because it requires data from at least two classes. In authentication, we cannot collect data from impostors because we cannot know who will be trying to gain unauthorized access. To solve this issue we use a hybrid system of CNN and SVM. We use a CNN model to extract features and these extracted features trains an SVM model—in one-class classification mode—that will perform classification.

We use GoogLeNet [23] CNN deep learning model because is the most efficient in terms of number of operations, accuracy and number of parameters [22]. Although, other models might perform better, the number of operations is significantly higher—longer operation time—and there is only a slightly improvement in accuracy. GoogLeNet is marks the gap between faster models—less accurate—and slower models—more accurate [22].

GoogLeNet has twenty two layers with parameters, twenty seven with pooling layers and an overall of one hundred layers that constitutes independent building blocks [23]. This work use this model until Loss3_Classifier layer—see Figure 13— and then feeds a support vector machine algorithm. Loss3_Classifier layer has 1000 outputs that represent

1000 features of the input image. In the GoogLeNet model, this layer feeds the grouping—classification—layer and usually is the starting point to fine tune the model—i.e. modification of the model for different applications. This work uses the Loss3_Classifier layer to get features that feeds a Support Vector Machine. SVM has the option—that deep learning does not have—to perform one-class classification, which is the required classifier for authentication.

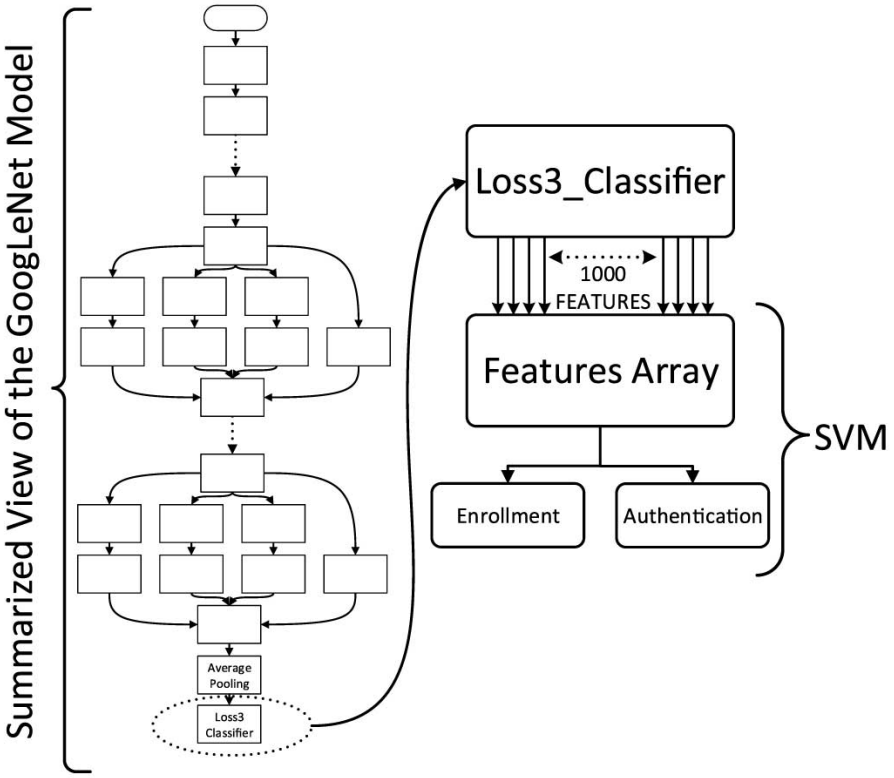

Figure 13. CNN-SVM hybrid model for authentication.

The array of features (12) stores all the extracted features of each heartbeat image. Each row of the array corresponds to a heartbeat and each column corresponds to a feature of a heartbeat image. In another words, vector (11) transposes from a row vector to a column vector and generates 1000 columns. Columns in array (12) has a fixed number of columns—1000 columns—because the deep learning model—GoogLeNet—always provides 1000 features at the Loss3_Classifier layer.

$$FA = \begin{bmatrix} ECG(img)_{HB_1,feat_1} & ECG(img)_{HB_1,feat_2} & \cdots & ECG(img)_{HB_1,feat_{999}} & ECG(img)_{HB_1,feat_{1000}} \\ ECG(img)_{HB_2,feat_1} & ECG(img)_{HB_2,feat_2} & & ECG(img)_{HB_2,feat_{999}} & ECG(img)_{HB_2,feat_{1000}} \\ \vdots & & \ddots & & \vdots \\ ECG(img)_{HB_{n-1},feat_1} & ECG(img)_{HB_{n-1},feat_2} & & ECG(img)_{HB_{n-1},feat_{999}} & ECG(img)_{HB_{n-1},feat_{1000}} \\ ECG(img)_{HB_n,feat_1} & ECG(img)_{HB_n,feat_2} & \cdots & ECG(img)_{HB_n,feat_{999}} & ECG(img)_{HB_n,feat_{1000}} \end{bmatrix} \quad (12)$$

Where, FA is Features Array, img is image, HB is heartbeat and $feat$ is feature.

4.1.5 Enrollment and Authentication

This work uses a Support Vector Machine algorithm to enroll—register a new user—and authenticate—validate a user—with the features obtained from the deep learning model—See Figure 13. We use SVM because supports one-class classification and outperforms other classifiers for ECG [32]. The non-linearity of Polynomial and Gaussian kernels provides a better fit to the data for one-class classification [134]. We use the Gaussian Kernel—also known as RBF: Radial Basis Function—because is more flexible than a polynomial curve. This flexibility allows to set closer boundaries to the training data, which leads to a better classification with most datasets [135]. Figure 14 shows the flexibility of the support vectors in the Gaussian kernel that fits the training data of the ECG. Figure 14 displays the general idea of the kernel behavior in a two dimensional graph; this work uses 1000 features in a model with 1000 dimensions that cannot be represented in a graph.

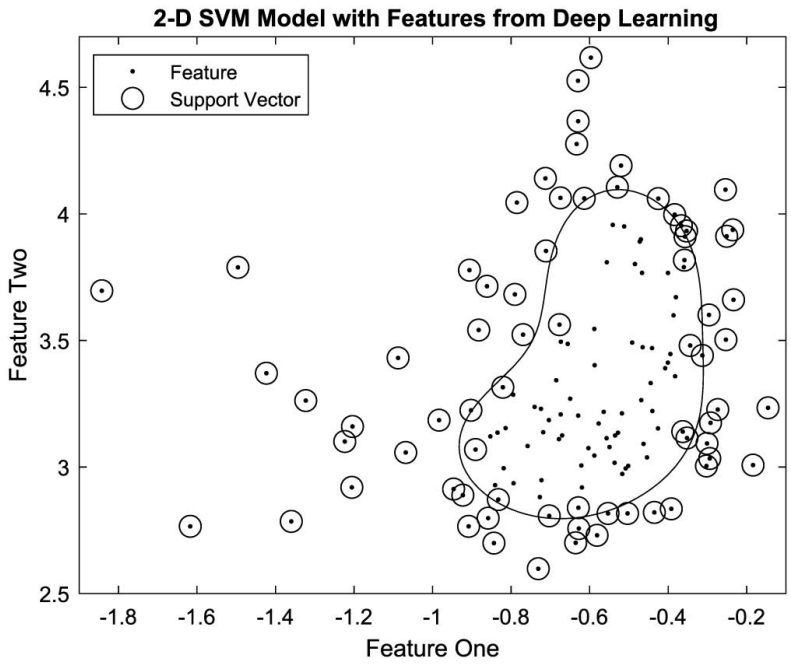

Figure 14. SVM One-Class boundary with two dimensions. The actual model in this work has 1000 dimensions.

Enrollment and Authentication use the array of features (12) to generate an SVM model—enrollment—or validate a genuine user—authentication. Enrollment is the process that registers a new genuine user in the biometric system and authentication is the process that determines if the user—trying to gain access—is genuine or impostor.

This thesis uses an ECG record of 120 seconds—2 minutes—for enrollment and 4 seconds for authentication. The enrollment ECG record is longer than our previous works [11], [41] because our preliminary tests shows that this model improves 160% the ERR if the ECG training record is 1 minute longer. A user perform enrollment only once, therefore this 1-minute increase in time does not significantly affects the user comfort. We maintain the authentication time in 4 seconds.

Enrollment uses the number of heartbeats available in an ECG record of 2 minutes long and authentication uses the number of heartbeats available in 4 seconds. The number of heartbeats available on each ECG record varies according to the heart rate of the subject during recording. This number will always be different because emotions and physical activities alter the heart rate of an individual. This change of heart rate determines the size

of the array of features (12), where the number of rows relates to the number of heartbeats available in the ECG record and—as mentioned earlier—the number of columns is constant—1000 columns—because of the number of features that the GoogLeNet model provides to the SVM.

The enrollment process uses the feature array FA—presented in (12)—to calculate the Lagrange multipliers α_i. A quadratic programming minimization function solves the minimization formulation of SVM and calculates the Lagrange multipliers α_i [136]. The enrollment process stores in memory or a database the Lagrange multipliers α_i—which is the SVM model— together with the feature array FA.

When a user attempts to authenticate, the deep learning model will generate a feature array FA_i—from a 4 seconds long ECG record. The authentication process uses this feature array FA_i, the Lagrange multipliers α_i and the enrollment template FA—generated from a 2 minutes ECG record—from memory or a database and calculates the classification function (13).

$$f(x) = \text{sgn}\left(\sum_{i=1}^{n} \alpha_i K(FA, FA_i) - \rho\right) \tag{13}$$

Where, $K(FA, FA_i)$ is the kernel function in terms of the Feature Array FA—enrolled user—and the feature array FA_i—user attempting to authenticate. As mentioned earlier, this work uses the Gaussian kernel (14).

$$K(FA, FA_i) = e^{-\|FA - FA_i\|^2} \tag{14}$$

The classification function (13) gives a positive or negative value that indicates if FA_i belongs to a genuine—positive—or an impostor user—negative. In SVM, the positive or

negative symbol is usually enough to classify the input data. The value that accompanies the symbol is a score that indicates how far is the input data from the support vector boundary established at the enrollment stage. This work uses this score value to add more flexibility to the SVM and applies a threshold to this score. The value of the threshold adjusts depending on the desired behavior of the classifier. Increasing the threshold value will make the authentication algorithm to reduce the number of wrongly accepted impostors—False Acceptance Rate—but will increase the number of wrongly rejected genuine users—False Rejection Rate. Decreasing the value will create the opposite effect; it will increase the number of accepted impostors and will decrease the number of wrongly rejected genuine users.

The number scores obtained from the classification function (13) are related to the number of heartbeats available in 4 seconds. As an example, if the feature array FA_i has three heartbeats, then we will have three scores. In order to reach a decision, only one value is necessary. This work calculates the average of these scores and applies the threshold to make a decision. We are calculating the average because all the scores are equally important and the obtained scores are close to each other.

4.2 Evaluation

This work evaluates the algorithm in two subsections. The first subsection compares it with related works that use different algorithms from deep learning. We use physionet [123] databases that were used in previous works [11], [41] to perform the comparison. From these databases, we use ECG records from 73 subjects from different age, gender and heart conditions. We extract 4 seconds of ECG for authentication and two minutes of ECG for enrollment.

The second subsection aims to compare the result with other works that use deep learning to perform authentication. The referenced works uses different databases to evaluate their work. But, all of them have one database in common which is the QT Database [126]. We use this database to compare our results with the referenced works.

We use the following hardware and software to run all the evaluations: Matlab 2019a running on a Windows 7 64 bits PC with an Intel Core i7 6700 3.40 GHz CPU, 16 GB RAM memory, an NVidia GeForce GTX 1060 6 GB GDDR5 GPU.

4.2.1 Comparison with previous related works

We test our algorithm using ECG data from our previous work [11] that has been evaluated against other ECG authentication algorithms. These algorithms do not use deep learning to perform authentication, but we present a comparison in this section to observe the performance of an algorithm using deep learning against algorithms that do not use it. Our previous work [11] has been proven to perform better than related algorithms; therefore, a comparison with this work implicitly compares it against the related works cited in [11].

We use 73 different ECG records—each record represents a user—from four Physionet databases: MIT-BIH Normal Sinus Rhythm Database [137], European ST-T Database [125], QT Database [126] and MIT-BIH Arrhythmia Database [124]. Each ECG record varies from 30 minutes to 24 hours and all these signals are part of different medical projects; this work does not discriminate signals that have heart conditions. From each record, we randomly chose four non-overlapping time locations to extract four sections of the ECG record. The first section is for enrollment and is 2 minutes long. The other three sections are for authentication and each one is 4 seconds long. Having this special ECG sections for enrollment and authentication, prevents the authentication process to use the same data for enrollment during the authentication of a genuine user. In another words, we authenticate a genuine user with different enrollment data of the same user.

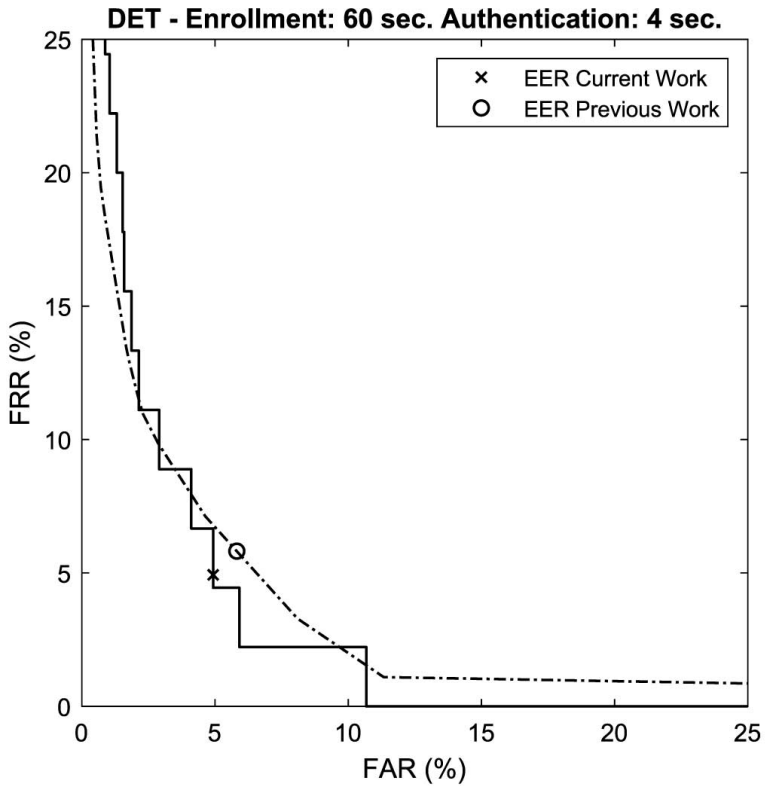

Figure 15. DET curve for the current Work and Previous Work. EER of current work is 4.9% and the EER of previous work is 5.8%

The evaluation follows the same procedure as in [11]. This procedure enrolls one user—genuine user—and performs authentication with all the users; this includes the genuine user. We repeat this procedure each time we change the genuine user. We finish the evaluation when all the 73 users have represented a genuine user and all the 73 users have represented an impostor. With this arrangement at any point of the evaluation, we had one genuine and 72 impostor.

Each authentication generates a score; we collect genuine and impostor authentication scores. With these scores, we select a threshold and apply the same threshold on the scores for genuine and impostors. The threshold on genuine scores will calculate the False Rejection Rate—FRR—and the threshold on impostor scores will calculate the False Acceptance Rate—FAR [39]. In order to perform a comparable evaluation with a previous work, we calculate the EER from the Detection Error Trade-off graph—DET—[39]. The DET graph displays the behavior of the biometric system when adjusting the threshold. It

displays the value of FAR against FRR when we adjust the threshold value [39]. The point at which FAR and FRR are equal is the EER [39] and this is the point that we use to compare this work with previous works.

Figure 15 shows DET curve of the results for this first experiment. These results show that our previous experiment has 5.8 % EER 5.8% and our current approach reaches 4.9 % EER. As mentioned earlier, this experiment runs under the same conditions and the same data as the previous work, Enrollment time of 60 seconds and authentication time of 30 seconds.

We perform another test on the same data, but we increase the enrollment time to 120 seconds—two minutes. Figure 16 shows the obtained results shows where we can observe that two minutes training lowers the EER to 2.84%.

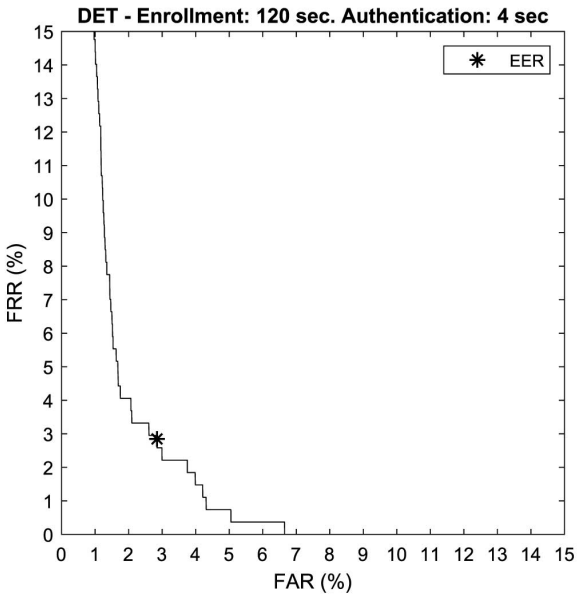

Figure 16. DET curve of current work. 2.84% EER with 120 seconds of enrolling time and 4 seconds of authentication time.

4.2.2 Comparison with related works

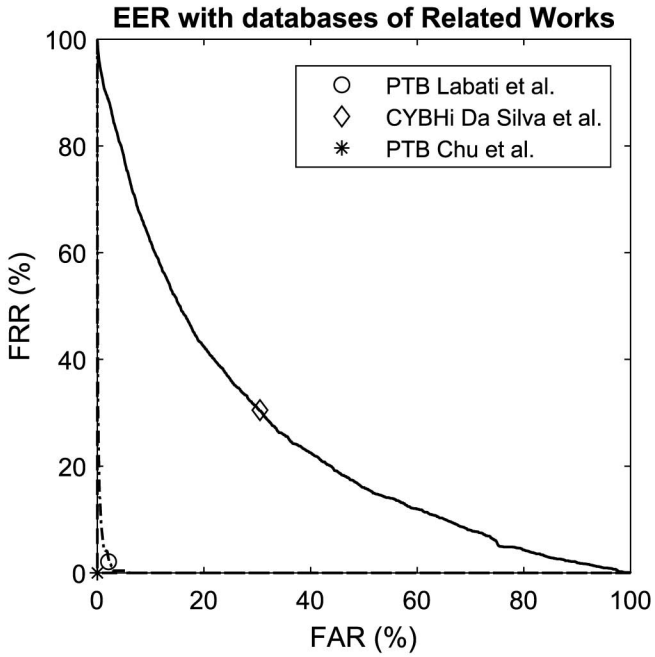

Figure 17. EER of this work with the databases used by related works. Labati et al. 2.1%, Da Silva et al. 30.5%, Chu et al. 0.01%

As mentioned earlier, many of the literature claim to perform authentication but they perform identification. This confusion is understandable in the deep learning field because

the main usage of deep learning is classification. Deep learning does not perform authentication—one-class classification problem. We found few studies that indirectly use deep learning in ECG authentication. In this section, we evaluate this work with the same databases as the related work. We follow the same procedures to the usage of the databases.

To compare this work with the results of Labati *et al*. [70], we use the same PTB database [71] from Physionet [123] to perform authentication. We follow their procedure before using the database. Their procedure removes ECG records from unhealthy users and removes noisy ECG records from 17 users. We use the database as indicated in their work and we run our testing to measure the performance of our algorithm. Figure 17 shows the result, where we get 2.1% EER.

Chu et al. [73] uses several databases—including the PTB database [71]—for their evaluation. Their algorithm combines two heartbeats in a vector. There is not enough number of heartbeats in the PTB database to complete the 1000 vectors; therefore, they use the same heartbeat in different vectors. The same heartbeat—in a combination with another—will be part of the authentication process once or more. We use they procedure with the database to try in our algorithm. Figure 17 shows our EER is 0.01% with the procedure they followed.

Da Silva Luz et al. [38] has a multibiometric approach but also present results in a unimodal approach. They use the CYBHi database [138], the data was collected on the same session and different sessions. We use the different session data to compare with their work because that is the closer scenario to a real application. To evaluate their work, they measure the noisy signals from the database and remove those that are a mean plus one and a half standard deviations out of their patron. We follow the same procedure and we obtained an EER of 30.5%, as we can see in Figure 17.

Table 1 presents a summary of the evaluation of this work with the same database and conditions of the related work. Same conditions refer to the time used for training and authentication and the removal of noisy signals. It is important to mention that most of

these works uses half of the dataset to train and the other half to test. In this work, the last row presents our results under our conditions. Our conditions do not remove any noisy signals and uses 2 minutes of ECG for enrollment time and 4 seconds of ECG for authentication time. The EER evaluation indicates the result that our algorithm has when using their same database and the same conditions. We also indicate the type of deep learning model they use and the databases that they have used.

TABLE 1. EER EVALUATION UNDER THE SAME CONDITIONS OF RELATED WORK

Algorithm	EER Evaluated	Model Used	Database
Chu *et al*. [73]	0.01%	ResNet – Own	PTB DB [71]
Labati *et al.* [70]	2.1%	CNN – Own	PTB DB [71]
Da Silva Luz *et al.* [72]	30.5%	CNN – Own	CYBHi DB [138]
Previous work [11]	4.9%	None	MIT-BIH N.S.R. [137]
This work	2.84% *	Hybrid – CNN-GoogLeNet & SVM	European ST-T [125] QT DB [126] MIT-BIH A. DB [124]

*2 minutes of ECG for enrolling and 4 seconds of ECG for authentication.

4.2.3 Discussion

In the first part of the experiments, the graphs show an improvement of almost 1% in EER of this work over our previous work. This is under the same enrollment time of 1 minute and authentication time of 4 seconds. When we increase the enrollment time to 2 minutes, the EER reduces to 2.84%. This is under the same authentication time of 4 seconds. The increased enrollment time represents a decrease of more than 2% in the EER. Increasing the enrollment time does not present a major inconvenient for a user. A user has to complete this process only when registering as a genuine user. The user does not perform this process often, unless is a new user or the biometric system lost user information. As an example, fingerprints is a fast biometric technology for authentication, but the enrollment time is the same or even more than ECG, that depends on the user that needs to place the fingerprint in the sensor for several times in different directions. We

expected this improvement with more enrollment time because the template will have more information that fits better to validate a genuine user or reject an impostor.

The second part of the experiments presents some challenges because all the related works use databases with different protocols. These protocols include the removal of noisy data, removal of records with health conditions or increasing the data by combining features. In order to perform a fair comparison we have to perform our test with the same protocols on the databases that they use. Each algorithm is build different and in order to use the diverse data, we have to do some adjustments in the algorithm to fit the data.

Table 1 summarises the results of this work with our previous work and related works that use deep learning. Chu *et al*. [73] reaches 0.59% EER. This is an excellent result for ECG authentication. We have two consider two things about their protocol. First the PTB databases has excellent ECG records—except few—that not contaminated with much noise. Second, the PTB database does not have enough data to generate one thousand vectors—each vector has two heartbeats—that their algorithm needs. Therefore, they use the same heartbeat in different vectors. They get an excellent result of 0.59% EER. We follow a similar procedure; however, our algorithm does not use vectors of heartbeats. We randomly duplicate heartbeats—as in their procedure—to add it to our testing data and treat them individually. We obtained an almost perfect result of 0.01% EER. However, we do not consider this as a valid evaluation. It might be a valid test for a proof of concept of the algorithm with vectors of two heartbeats. However, this is not feasible in an application because during authentication we cannot extract the same heartbeat two or more times. A user always provides different heartbeats, they are similar but not equal. Duplicating the heartbeat can increase the chances of accepting genuine users but also increase the chances of accepting impostor. The opposite can happen, increases the chance of rejecting impostors but increase the chance of rejecting genuine users. Laboratory tests does not reflect this changes because the test counts each duplicated heartbeat as a new heartbeat, but in reality they are not.

Labati *et al.* [70] also uses the PTB database. They use the ECG records of health patients only and removed records of 17 patients that were noisy. We follow the same

procedure and remove records of the 17 patients that we consider them noisy. Our result of 2.1% EER is close to their result of 3.37% EER. On of the reason for the difference is that we use a pre-trained model with millions of images. They train a CNN model with 771 hours of ECG raw signal. It is enough data to train a CNN model but the pre-trained model has 5 times more data, which reflects in the results. Another aspect to consider is that they use three leads—multi-sensor fusing—while we tested with only one lead. The different removed signals might have an impact on the results, but we have to consider the fact that we use only one lead instead of three.

Da Silva Luz *et al.* [72] use their own collected database and made it public [138]. We follow the same procedure to test the algorithm with the database collected in two different sessions. We obtained an EER of 30.5%, which is close to their 26.58% EER. The high EER rate of our algorithm and their algorithm is because this database is highly contaminated with noise. The filtration techniques that we use are not enough to filter them. To alleviate the effects of noise, they remove noisy signals that has one mean plus one-and-half standard deviation away from their patron. We follow the same procedure but we cannot guarantee that we remove the same noisy signals. This difference on noisy signals might have an impact on the results. However, it is important to mention that we are using a pre-trained model for image classification. This has the advantage on saving resources. We do not need to train our model with an extensive dataset and we do not need high-end GPU hardware to perform our training. In addition, pre-trained models keep getting better with more data and improved architectures. With our work, we can upgrade the pre-trained model to a better one and this will improve the results.

Using pre-trained CNN models improves the results because the feature extractor has an outstanding performance, especially if the pre-trained model is the result of millions of images. The automatic extraction of features helps with this issue in ECG authentication. ECG as images has proven to have acceptable results that are comparable with similar works, we the advantage that we can improve with any pre-trained model that becomes available.

Chapter 5.

ECG Biometric Fusion with Fingerprint

This chapter describes the ECG and fingerprint fusion mechanism for the bimodal authentication algorithm. Bimodal fusion takes the speed and accuracy of fingerprints and the security of ECG and combines it in solution that has the best of both. In the next sections of this chaper we present the design and the evaluation of the algorithm.

5.1 Design of Bimodal ECG – Fingerprint Authentication Algorithm

The Bi-modal authentication algorithm is the biometric fusion of the independent results (decision level fusion) of Fingerprint and ECG unibiometric algorithms.

The fingerprint biometric uses the MINDTCT algorithm as the minutiae extractor and the BOZORTH3 algorithm as the minutiae matcher [65]. The latter algorithm will output a score reflecting the match level between the input and stored template(s). We explained more details about the algorithm in section 2.5.

The bimodal authentication algorithm uses two independent unibiometric results and fuses them to reach a final decision on authentication (Figure 18). We employed a decision-level fusion scheme because the matchers of unibiometric methods produce two types of results: a binary output for ECG and a score for the Fingerprint. As we explained in Section 3.1, the ECG matcher uses an SVM classifier to perform authentication. Rather than calculating a match score; the SVM classifier produces a binary result: a match or a non-match. The Fingerprint matcher returns a score that measures the similarity between the stored and input templates. To fuse at the score level, both matchers need to provide score values. However, because ECG-SVM only provides a binary value, we could not combine these results at the score level. Instead, these results could only be fused at the decision level, which requires compatible results from the matchers. Hence, we converted the fingerprint score value into binary results using a threshold (any score above the threshold, is a match; otherwise, it is a non-match).

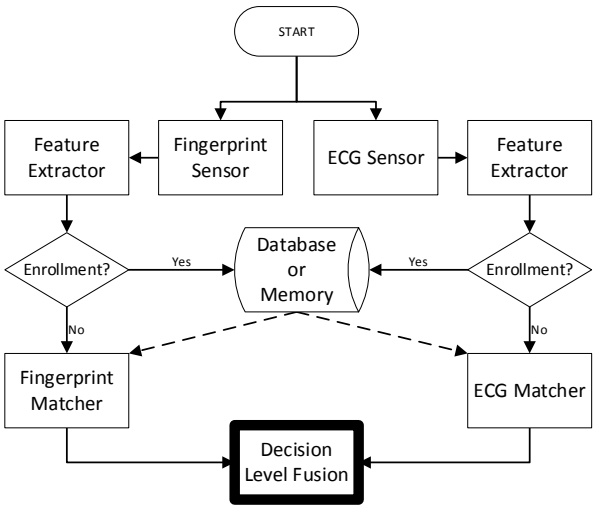

Figure 18. Bimodal Authentication Algorithm

Figure 19 illustrates the mechanism of decision-level fusion, which can achieved using two alternative approaches:

- Fusion Method A: Use the fingerprint results first and the ECG results second.
- Fusion Method B: Use the ECG results first and the fingerprint results second.

We evaluate these two approaches in Section 5.2.2 to identify which one produces the lowest EER.

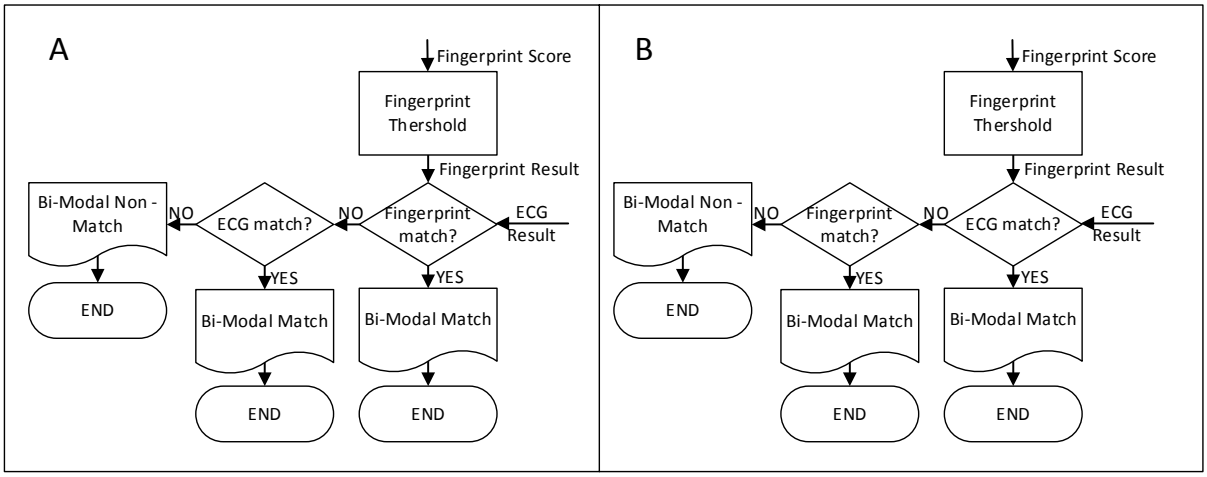

Figure 19. Bimodal Decision Level Fusion. a) Fusion Method A. b) Fusion Method B.

Figure 19 represents our approach. We adopt this approach because other decision level fusion schemes are not applicable for the proposed algorithm. For instance, the majority votes technique is unfeasible since we only have two classes (i.e. genuine and impostor) with two matchers. In the case of a tie, the algorithm would not be able to render a final decision. A linear weighted fusion scheme is also not applicable, as we need to assign a weight for each classifier that corresponds to its performance. To measure the performance we need to calculate weights based on the results of the fusion. Section 2.7.2 describes that weight calculation is the major drawback on this scheme. In this study, we do not have a suitable number of matchers (i.e., ECG-SVM and Fingerprint) or classes (i.e., match or non-match) to calculate an appropriate weight. Similarly, Behavior Knowledge Space requires a large number of classes and datasets to work properly. In this work we have two classes (match or non – match), this is a limitation in applying Behavior Knowledge Space.

5.2 Evaluation of Bi-modal authentication algorithm

We evaluate the proposed algorithm in three stages. In the first stage, we evaluate the MINDTCT minutiae extractor and BOZORTH3 fingerprint matching algorithm [65] with the same fingerprint images used to evaluate the bimodal authentication algorithm. In the second stage, we identify the best fusion method for the bimodal authentication. Finally, in the third stage, we compare our algorithm to existing approaches.

Biometric authentication systems have several operating points pertaining to the threshold(s) of the biometric matcher. Each operating point represents a trade-off between errors; an operating point that decreases the FAR also typically reduces the TAR. A low TAR results in a higher False Rejection Rate (FRR) ($FRR = 1 - TAR$) as more genuine users are rejected. Conversely, if the operating point produces a lower FRR, less genuine users will be rejected, but more impostors will be accepted. An effective biometric system should minimize the FAR and FRR. DET graphs displays the relationship between FAR and FRR to allow us to choose the best operating point [42]. Hence, to compare biometric matchers, we can find the DET point where FAR is equal to FRR. We call this DET point the Equal Error Rate (EER). Hence, in this section, whenever possible, we will use the EER metric to assess the effectiveness of the biometric matcher.

5.2.1 Fingerprint Evaluation

We used fingerprint images from 73 subjects in the DB1 category of the FVC2006 database [139]. Each subject has 7 images, consisting of 1 image for enrollment and 6 images for authentication. These images were extracted using an electric field sensor (AuthenTec) [139]. The size of each image is 96 x 96 pixels with a resolution of 250 dots per inch (dpi). The purpose of this evaluation was to assess the performance of the fingerprint authentication algorithm. Later in Section 5.2.2, we combine this data with ECG data to evaluate the bimodal authentication algorithm. We will use the results of this evaluation to compare the performance of the state of the art unibiometric fingerprint method with the proposed bimodal biometric technique.

We ran the evaluation in a batch mode, enrolling the subjects one by one and evaluating them against all non-enrolled subjects. Therefore, each time we enroll a subject, we evaluate their fingerprint image against that of 1 genuine subject and 72 impostors. We repeated this process with each new subject until we complete the enrollment of all subjects. Each evaluation yielded a score and we applied a threshold to the score to obtain a matching decision. We evaluate a range of thresholds that goes from 0 to 100 in steps of 1. We plotted these results in a DET graph (Figure 20).

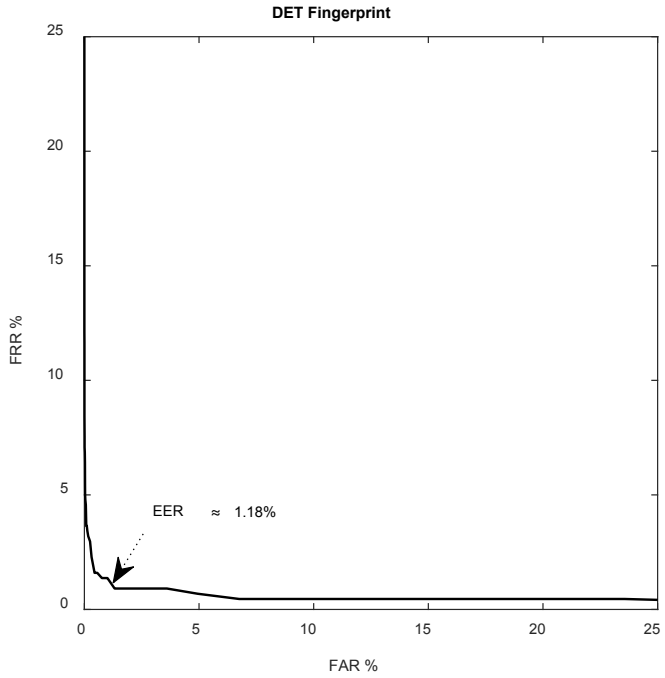

Figure 20. Direct Error Trade-off (DET) Graph For Fingerprint Performance.

The evaluation of the unibiometric fingerprint system produces a FAR of 1.05% and FRR of 1.37%. We use the approach presented by [140] to obtain an approximate EER of 1.18%.

5.2.2 Bimodal Algorithm Evaluation

To evaluate our bimodal biometric algorithm, we used the dataset previously described in Section 5.2.1 for the fingerprint images and the dataset previously described in Section 3.2 for the ECG records. We combine an ECG record with a fingerprint image to generate a "virtual" subject for the bimodal authentication evaluation. Therefore, our final database

consisted of 73 subjects, where each subject had 7 ECG records and 7 fingerprint images. We use one fingerprint image with one ECG record (60 seconds long) for enrollment and the other 6 ECG records (4 seconds long) with the corresponding 6 fingerprint images for authentication.

We performed the evaluations with Matlab 2016a running on a Windows 7 64 bits PC with an Intel Core i7 CPU of 2.8 GHz, 8 GB RAM memory. We ran our experiments in a batch mode; we enrolled (with one 60-second ECG record and one fingerprint image) all the subjects one by one and evaluated them against the non-enrolled subjects. Therefore, each time we enrolled only one subject we evaluated our algorithm against 1 genuine subject and 72 impostors. We repeated the process with each subject until we enrolled them all.

As previously defined in Section 5.1, we evaluated two fusion schemes, fusion Method A and fusion Method B. These schemes differed by the type of matcher (fingerprint and ECG) that was used first. As illustrated in Figure 19, we first evaluated Fusion Method A and second we evaluated Fusion Method B. Because ECG uses SVM, then a threshold is only applicable to the fingerprint method. We evaluated the bimodal authentication algorithm for several thresholds (ranging from 0 to 100 in steps of 1) and we plotted a DET graph (Figure 21) with the obtained results.

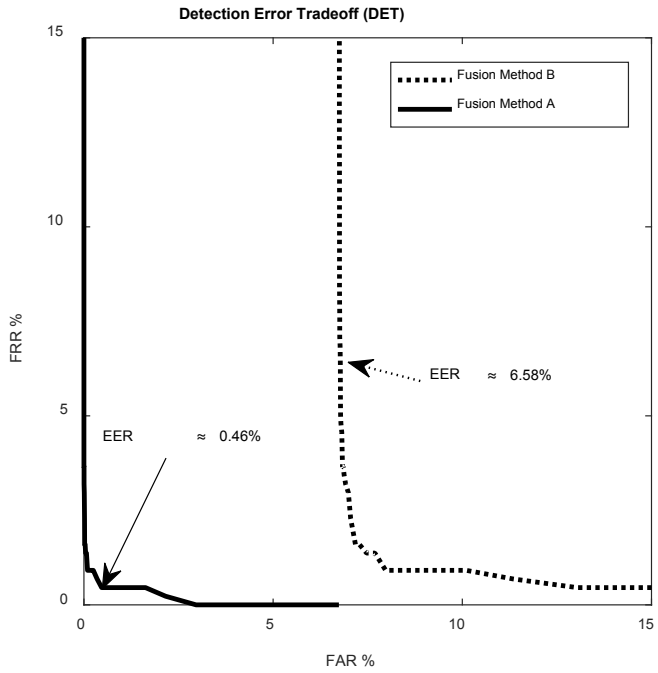

Figure 21. DET Graph For Bimodal Fusion Method A and Method B

In Figure 21, we see that fusion method A performs better than fusion method B. Fusion method A displayed a FAR of 0.47% with a FRR of 0.46%, which give us an approximate EER value of 0.46%. Method B displayed a FAR of 6.78% with a FRR of 6.39%, which give us an approximate EER value of 6.58%. The difference between fusion method A and fusion method B is around 6% in terms of EER. We conclude that fusion method A is the better mechanism for bimodal authentication using ECG and fingerprint.

Fusion method A produced better results than fusion method B since the fingerprint algorithm uses a threshold on the score, which allows us to maximize the TAR at the expense of the FAR. Then, the ECG algorithm decreases the FAR, which results in a lower EER. This is not the case when the ECG matcher executes before that of the fingerprint. The ECG matcher does not have a threshold to adjust and hence the TAR cannot be maximized by moving the threshold.

The bimodal approach is more effective than the fingerprint or ECG biometric methods alone. In section 5.2.1 we found an EER of 1.18% for the fingerprint unibiometric scheme. Fusion method A had an EER of 0.46%; this shows an improvement of 0.72% for the EER. These results showed that our bimodal biometric method performs 2.5 times better than the

fingerprint unibiometric approach. Moreover, in Section 3.2 we show that ECG-SVM has a FRR of 0% ($FRR = 100 - TAR$) with a FAR of 7%. We cannot generate a DET graph for the ECG-SVM unibiometric scheme because the SVM matcher returns a decision (match or non-match) and not a score. Therefore, we examine the FAR of our bimodal authentication method when the FRR is 0%. To do this, we use the DET graph of the fusion method A (see Figure 21) and we obtain a FRR of 0% when FAR is 2.96%. These results show that our bimodal method also performs better than the unibiometric ECG-SVM unibiometric scheme.

5.2.3 Evaluation with Existing Works

In this section, we compared our proposed scheme to those of [84] and [15]. However, [84] and [15] used different databases to evaluate their algorithms. For [84], they captured their own ECG and fingerprint dataset. This dataset is not publicly available and the number of testing subjects is not specified. Sing et al. [15] used ECG data from Physionet and fingerprint scores from NIST-BSSR1 [86]. In our work we used ECG data from the same source as Sing et al. [15]; however, for fingerprint we used the FVC2006 database [139].

Reported evaluation conditions and datasets are different among all three algorithms; therefore, to achieve a proper evaluation, we implement the other two algorithms and evaluate them with the same datasets.

Work [84] described their results in terms of FRR (0%) and FAR (2.5%). The dashed line in Figure 22 shows the DET curve of their algorithm with the dataset we employ to evaluate our work. We can observe in Figure 22 that the EER was approximately 25%. Among the many factors a high EER value, one of them is normalization. Because Manjunathswamy et al. work did not normalize the ECG signal; changes in the heart rate would have affected the matcher results. In contrast, our data is composed of users with different heart rates, which affects the response of their algorithm.

The dotted line in Figure 22 shows the DET curve for the evaluation of multimodal algorithm by Sing et al. [15]. Similarly, to evaluate this algorithm we used the same datasets used in our algorithm. While the authors reported an EER of 1.52% for their algorithm, our evaluation with our data revealed that their algorithm has an EER closer to 12%. The number of features and length of ECG records are the reason for the discrepancy between these results. They extracted 20 features from the ECG signal. The extraction of some of these features is not always possible. Filtration of the noise in the ECG signal can render some fiducial points very difficult to locate; therefore, it will prevent to extract a feature. If a feature is missed, then the whole heart beat is discarded. This loss of information causes the EER to be higher. Another aspect that affects the EER is the length of the ECG records that they used. The minimum length they used for enrollment and authentication is 3 minutes; some records can be as long as 12 hours for enrollment and 12 hours for authentication. In contrast, our dataset consists of 60 seconds for enrollment and 4 seconds for authentication.

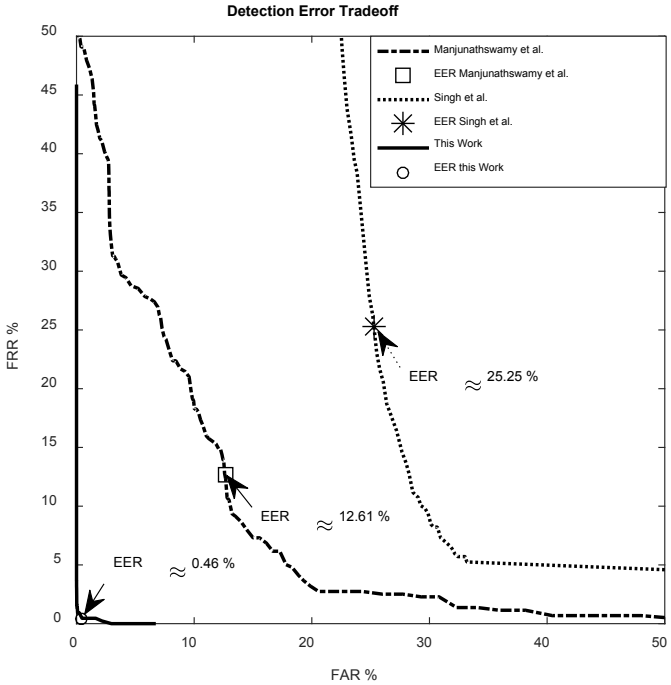

Figure 22. Multimodal DET Graph: Related Works Comparison.

Table 2 compares the characteristics of our work with that of the previously cited works. To compare fairly and accurately the results of all three studies, our evaluation used

the same data and parameter sets. Table 2 presents the results reported by these works and the results from our evaluation of the algorithms. We do this to present a fair comparison. We used the following parameters for all three compared methods: Enrollment time of 60 seconds; authentication time of 4 seconds; number of Subjects was set to 73, fingerprint data from the FCV2006 database [139], and ECG data from the Physionet database. The physionet database contains several databases. From those we use the European ST-T Database, the MIT-BIH Normal Sinus Rhythm Data-base, the MIT-BIH Arrhythmia Database and the QT Database [85]. Our proposed bimodal method has an EER of 0.46%, which is lower than Sing et al. [15] (EER of 12.61%) and Manjunathswamy et al. [84] work (EER of 25.25 %).

TABLE 2. COMPARISON OF MULTIMODAL RESULTS

	Manjunathswamy et al. [84]	**Singh et al.**[15]	**This Proposed Work**
Number of Features	ECG: 11 features FP: 2 set minutiae -1 ridge endings -1 bifurcations	ECG: 20 features FP: 1 set of minutiae (ends and bifurcations)	ECG: 8 features FP: 1 set of minutiae (ends and bifurcations)
Level of Fusion	Feature level	Score level	Decision level
Reported Results	FRR: 0 % FAR: 2.5 %	EER: 1.52 %	EER: 0.46 %
Results with Same Parameters	EER: 25.25 %	EER: 12.61%	EER: 0.46 %

Furthermore, our approach uses decision level in contrast to the approach by Sing et al. [15] which fuses with a weight sum rule at the score level. Fusion at the decision level provides independence between the matchers (i.e., each matcher works as a unibiometric until fusion). An independent matcher is the one that generates a score and makes the matching decision. When fusing at the score level, matchers are not independent because each matcher generates a score and another decision module fuses all these scores and makes a matching

decision. When fusing at the decision level, matchers are independent and another decision module provides a final decision (match or non-match) based on the decision results of the unibiometrics matchers. Prabhakar et al. [80] found that, in a multibiometric approach, independent matchers perform better; therefore, that is one of the reasons that decision-level fusion improves the performance in our multibiometric approach.

Chapter 6.

R-Peak detector

In ECG authentication is important to have a low Average Time Error in the R-peak detection algorithm. A low average time error means a more accurate location of the R peaks that will reduce errors with ECG authentication. This chapter describes an R-peak detection algorithm with a low average time error as a signal-processing tool.

6.1 Design of the R-Peak Detection Algorithm Based on Differentiation

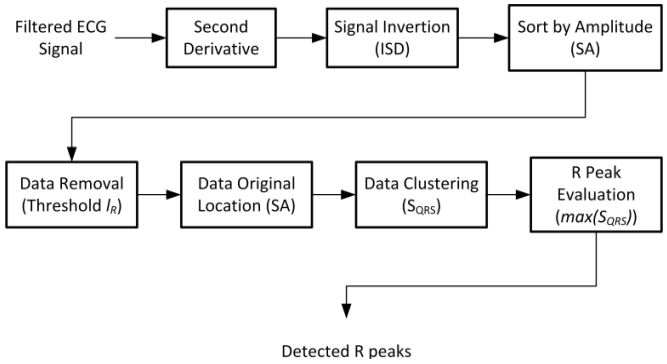

Figure 23. Diagram of the R peak detection logic

Figure 23 shows the various processing stages of the proposed R peak detection algorithm. The first stage calculates the second derivative of an ECG signal (ECG signal shown in Figure 24a and the second derivative is shown in Figure 24b). Our R peak detection algorithm takes the time series produced by the second derivative and inverts it as shown in Figure 24c. This produces the Inverted Second Derivative (ISD) record.

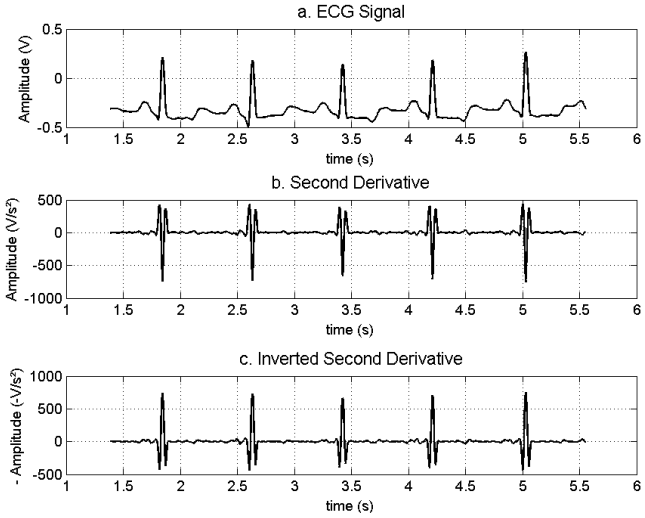

Figure 24. Stages of R-Peak detection. a. Segment of ECG record 100 from MITDB Arrythmia Database. b. Second Derivative of the filtered ECG record. c. Inverted of the second derivative, where the detection of the R peak will start.

The next stage sorts all the values in the ISD series in a descending order while maintaining in a data structure called SA the timestamp corresponding to each value of the ISD series as shown in (15). This will place at the beginning of the SA, all the values that belong to a QRS complex. See Figure 25a.

$$SA = \begin{bmatrix} a_0 & a_1 & \cdots & a_{f-1} & a_f \\ t_{a0} & t_{a1} & \cdots & t_{af-1} & t_{af} \end{bmatrix} \qquad (15)$$

In (15), a is a value in the ISD series, t is a timestamp that corresponds to value a and f is the length of the signal. In the sorted data structure SA (Figure 25a), it is certain that the first set of values belong to QRS complexes in the ECG record. From this set, the exact quantity of values that corresponds to QRS complexes is determined in the next step. To determine the values that are part of the QRS, this work introduces the variable l_R as the index of the last

possible QRS related value in the data structure shown in (15). Equation (16) calculates l_R and considers the maximum heart rate of a human being in variable HR_{max} which has beats/second as units. It also uses the time length of the signal (t_f in seconds) and the number of samples that are part a QRS complex (S_{QRS} in samples/QRS). The foundation of this concept is that an R peak is a single sample; but the whole QRS complex (that includes the R-Peak) is represented by several samples.

$$l_R = HR_{max} \times t_f \times S_{QRS} \qquad (16)$$

The value of HR_{max} is constant and is approximated at 220 beats/minute (~3.66 beats/second). This approximation of the HR_{max} is based on the formula introduced in [141]

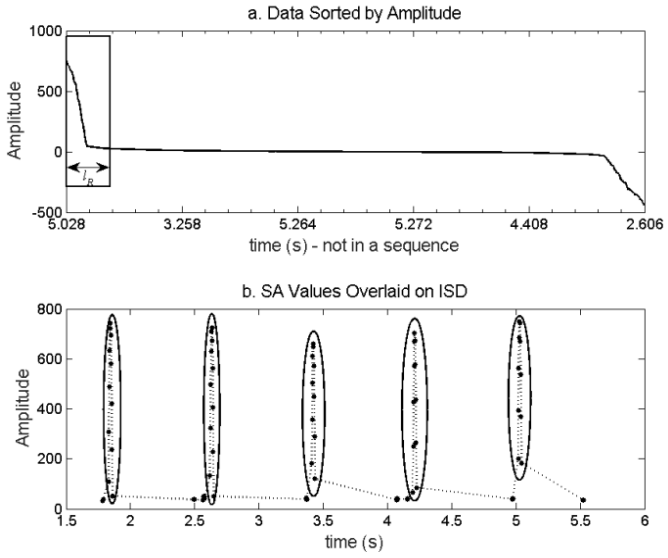

Figure 25. SA Data Structure

The number of samples in a QRS complex (S_{QRS}) is proportional to the sampling frequency. The higher the sampling frequency, the more samples per QRS complex we will have. Several experiments calculate the constant of proportionality k at various sampling frequencies. The experiments measure the number of samples that creates a QRS complex at a certain frequency and divides the average of the number of samples by the average of the sampling frequency in order to obtain k. The value of the calculated constant of proportionality k is 0.019843. Equation (17) shows the resultant expression to calculate S_{QRS};

where, f_s is the sampling frequency. The final calculation rounds S_{QRS} to the nearest integer value.

$$S_{QRS} = k \times f_s\,, \qquad S_R \in \mathbb{Z} \qquad (17)$$

From the data structure of (15), we eliminate all the entries beyond the index l_R. This means that most of the values that do not belong to QRS structures are removed. We overlay the values of SA to their original location in the ISD (by relying on the maintained timestamps to do so) as shown in Figure 25b. This produces clusters of overlaid values that mostly represent QRS complexes, with few exceptions that are easy to spot at this point. We consider a set of points to belong to the same cluster if they respect a continuity test. This test stipulates that two consecutive points belong to the same cluster if and only if the difference between their timestamps is equal to Δt (where $\Delta t = 1/f_s$). At the end, we will have several clusters with different number of samples in each one of them. If the number of samples in a cluster is less than S_{QRS}, then the cluster is discarded as it is assumed to be a non QRS complex. S_{QRS} is the minimum number of samples that a cluster requires in order to be considered as a QRS. This means that each QRS cluster does not necessarily has to have the same number of samples; each cluster can have different number of samples, but the minimum number of samples required in a cluster is S_{QRS}.

The last step of the algorithm is to detect exactly which one of the clustered amplitudes corresponds to an R peak. This is done by finding, for each cluster, the corresponding values in the original ECG record whose maximum will constitute the R peak in that cluster.

Section 6.2 presents the evaluation of this R-peak detection algorithm. The low average time error of this algorithm is an important part to guarantee accurate results for ECG authentication. Another important part of any biometric system is the evaluation. Evaluation of biometrics system regularly uses the EER parameter obtained from the DET graph. A multimodal biometric approach can generates data with multiple thresholds. The next section presents the second contribution of this work, which is a method to calculate EER from a DET graph in a multi-threshold biometric system.

6.2 Evaluation

This section evaluates the R peak detection algorithm that was developed to detect the R peaks in order to extract the biometric features to implement a biometric authentication with ECG.

6.2.1 Experiment Setup

The algorithm was tested with 48 ECG records from the MIT-BIH Arrhythmia Database [124]. Each record has a duration of 30 minutes and they were obtained from 47 individuals. The sampling frequency for all the records is 360 samples per second and with a resolution of 11 bits in a 10mV range. Each record has been manually annotated in the databased by cardiologists; they indicate the different characteristics that the signal has, including the locations of R peaks.

The ECG database was downloaded from Physionet [85], each download file includes the ECG data and annotations. The information was extracted and processed with Matlab software and the Matlab tool for Physionet [142].

To evaluate our algorithm we are using as reference the location information of an R peak that is provided by the database. The Physionet tool for Matlab [142] performs the calculations of sensitivity, positive predictivity and average time records according to standardized norms [143]. The tool takes as patron of comparison, the annotation and data files from Physionet; and as test data, the information that we provide from our algorithm.

Among all the annotations in the ECG records, some of them indicate that the signal is unreadable at a particular segment of the record. These segments have been removed from the testing and only the annotations that indicate the presence of an R peak have been considered in the experiment.

6.2.2 Results

The obtained results from our algorithm, along with the results for Method V are presented in Table 3. It can be observed that our algorithm has good positive predictivity and sensitivity rates, while maintaining a low average time error (compared to Method V). This is due the use of a threshold based on the sampling frequency rather than amplitude and it is fixed for the entire signal. The use of the second derivative helps on maintaining a low average time error. The use of a second derivative introduces a phase shift of the signal that is always equal to $2\Delta t$. Since this is a fixed value, the time location of an R peak can be corrected by removing this phase shift of $2\Delta t$.

TABLE 3. RESULTS FOR R-PEAK DETECTION ALGORITHM

Algorithm	FN	FP	TP	Se (%)	+P (%)	Average Time Error (ms)
Proposed	257	351	44870	99.430	99.224	4.9535
Benitez et al. [110] ***with Second Derivative (Method V)***	2112	884	107344	98.07	99.18	6.50

Chapter 7.

Multi-threshold Evaluation Method

Chapter 5 presents a bi-modal authentication algorithm. This algorithm fuses at the decision level and has multiple thresholds in each matcher. In order to evaluate a multiple threshold approach we have develop an evaluation method for that adjust multiple thresholds to a single threshold evaluation method. This chapter describes the approach that we have proposed to determine the DET graph with more than one threshold and the mathematical method to calculate the EER for non-normal distributions.

7.1 EER Calculation and DET Approximation in a Multi-Threshold Biometric System

The proposed solution for EER Calculation and DET Approximation in a Multi-Threshold Biometric System is a combination of two approaches. First, we use an intersection sensitive algorithm [144] for calculation of intersection of curves and apply it for the EER calculation in the DET curve. We cannot directly apply this approach in a multi-threshold biometrics because the DET is not a single curve. Second, in order to apply the DET calculation in a

multi-threshold biometric, we determine the DET by using a hull algorithm and then we calculate the EER with the algorithm for curves intersection.

7.1.1 Calculation of EER by intersection of curves

We calculate the EER by determining the intersection point of the DET curve with the line $x - y = 0$. This line has a slope $m = 1$ or an inclination of 45 degrees. We can rewrite the equation as $x = y$. The DET curve has FAR values in the x axis and the FRR values in the y axis; therefore we can say that the line $x = y$ represents the points where FAR and FRR are equal. The intersection of this line with the DET curve indicates the location of the EER. We depict the EER calculation in Figure 26.

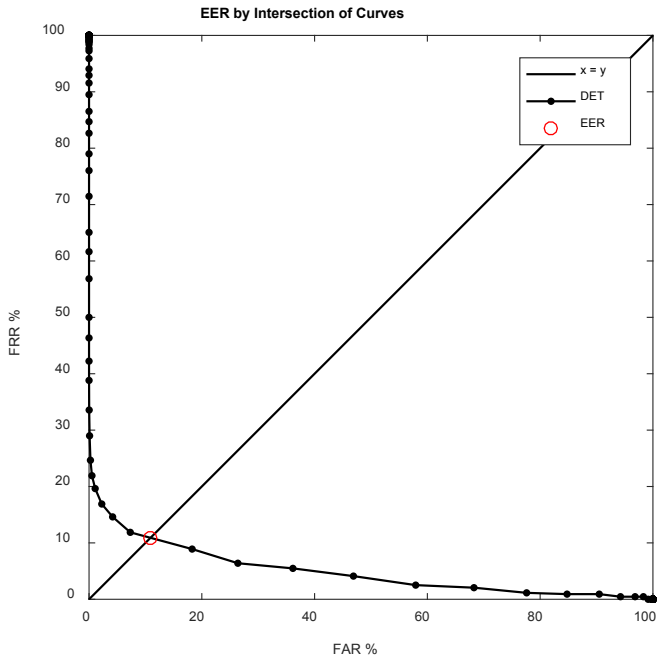

Figure 26. Calculation of EER by intersection of Curves

We use computational geometry to find the intersection points. In computational geometry there is not continuous data, we have samples that mimics continuous data, but it is unlikely that we will find a sample where DET and $x = y$ intersects. In Figure 26 we can see that no sample (dot) from DET intersects the line $x = y$. This line intersects a line

between two samples. To solve this issue, we use an algorithm known as the *intersection sensitive* [144]. This algorithm creates segments between each sample. These segments represent a line. We would not find a sample that intersects the two curves, but we will find a segment that intersects two curves. These segments represent a line and we use basic line equation in order to find the intersection point. We have to do that for every segment on the curves. The complexity to run this algorithm is $O(n^2)$. Literature presents a solution to reduce the complexity. The solution is to use a *plane sweep algorithm* [144]. This algorithm does a swipe through all the segments and marks the segments that have a possible intersection. With this information, we only process segments with possible intersections. This approach has a complexity $O(n \log n)$.

We calculate the EER as the intersection of DET and the line $x - y = 0$ as follows:

$$DET = \left[P_{0(x,y)}, P_{1(x,y)}, \dots, P_{n-1(x,y)}\right] \qquad (18)$$

Where P is a sample that forms the DET curve and x, y corresponds to the coordinates of the DET sample. From (18) we subtract the coordinates $x - y$ for all the samples and stores in a vector $A_{(i)}$.

$$A_{(i)} = P_{i(y)} - P_{i(x)} \qquad 0 \leq i < n \qquad (19)$$

The subtraction in (19) gives us a positive or negative value for a specific point $A_{(i)}$. Two consecutive points ($P_{(i)}$ and $P_{(i+1)}$) forms a segment. When two points have different sign, it means that they are located in different sides of the line $x - y = 0$; therefore, that segment forms a line that intersects $x - y = 0$.

We store the values of different sign in a vector $B_{(i)}$ as ones or zeros (see (20)). We identify with "ones" the first point of a segment that intersects the line $x - y = 0$. There is a special case where the sample point coincides with the line $x - y = 0$; if that is the case, we also mark it with a "one".

$$B_{(i)} = \begin{cases} 1, & sgn(A_{(i)}) \neq sgn(A_{(i+1)}) \\ 1, & sgn(A_{(i)}) = 0 \ \vee \ sgn(A_{(i+1)}) = 0, \\ 0, & otherwise \end{cases} \qquad 0 \leq i < n \tag{20}$$

Where,

$$sgn(x) = \begin{cases} -1, & x < 0 \\ 0, & x = 0 \\ 1, & x > 0 \end{cases} \tag{21}$$

We determine the index of the first point of the segment that intersects $x - y = 0$ and store it in t according to (22).

$$t = i, \ \text{only if} \ \ B_{(i)} = 1 \tag{22}$$

We find the EER by using the segment that we know it intersects the line $x - y = 0$. We know that the line equation is:

$$y = mx + c \tag{23}$$

Where the slope m is,

$$m = \frac{y_2 - y_1}{x_2 - x_1} \tag{24}$$

We adjust the slope m from (24) to the points of DET and the index t

$$m = \frac{P_{t+1(y)} - P_{t(y)}}{P_{t+1(x)} - P_{t(x)}} \tag{25}$$

From (23) we can calculate c according to (26) and (27).

$$c = y - \frac{P_{t+1(y)} - P_{t(y)}}{P_{t+1(x)} - P_{t(x)}} x \tag{26}$$

$$c = P_{t(y)} - \frac{P_{t+1(y)} - P_{t(y)}}{P_{t+1(x)} - P_{t(x)}} P_{t(x)} \tag{27}$$

EER will be located at some point in line $x - y = 0$; therefore, we can say $y = x = EER$. We can replace eq. (23) as:

$$EER = mEER + c \tag{28}$$

$$EER = \frac{c}{1 - m} \tag{29}$$

In (29) we can replace terms with (27) and (25).

$$EER = \frac{P_{t(y)} - \frac{P_{t+1(y)} - P_{t(y)}}{P_{t+1(x)} - P_{t(x)}} P_{t(x)}}{1 - \frac{P_{t+1(y)} - P_{t(y)}}{P_{t+1(x)} - P_{t(x)}}} \tag{30}$$

We present our EER calculation method in (30). This approach can be applicable in normal and non-normal data distribution. In Biometrics the data distribution usually is not normal.

7.1.2 DET curve and EER for multi-threshold biometrics

In order to estimate the EER where two or more thresholds are involved, first we need to plot the DET curve. Plotting the DET curve with more than two thresholds gives us several points; therefore, we need to apply a different approach for the estimation of EER.

With the points in the graph that represent a FAR with its respective FRR, we group all of them in a convex hull. With this hull we use the lower side of the convex hull as the DET graph. In this hull, we estimate the value where FAR is equal to FRR; this point is known as the Equal Error Rate (EER).

The convex hull algorithm that we used is the Quick Hull Algorithm [145]. This algorithm uses a similar approach as a quick sort algorithm. It divides the data in order to operate. First, we find the maximum and minimum points along the x-axis and trace the main line between these two points, as depicted in Figure 27a. This line divides all the points in two groups. In each group, we calculate the distance of the points to the line. From each group we get the points with the maximum distance and we trace a triangle with the main line as the base (see Figure 27b). All the points that are inside these triangles are ignored and we work only with the points outside the triangle. Next step is to calculate the points with the maximum distance with respect of the sides of the traced triangles (see Figure 27c). We repeat this process until there is no more points left. The final hull is presented in Figure 27d.

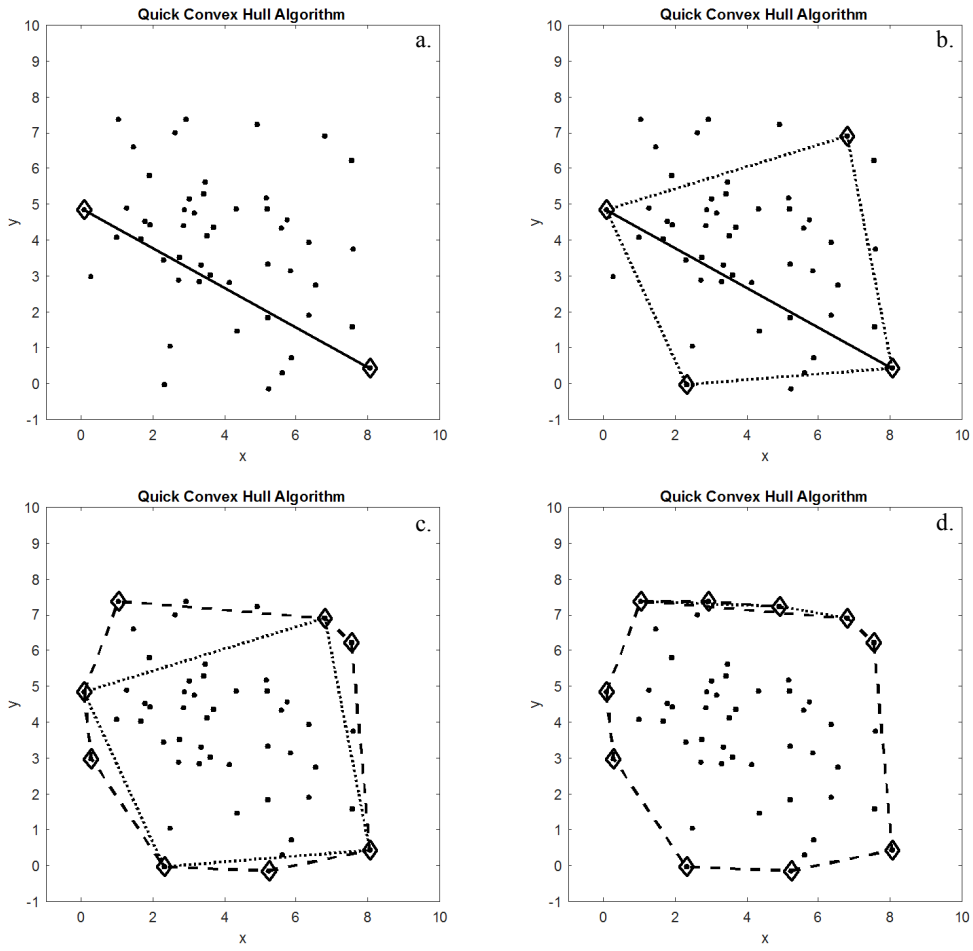

Figure 27. Steps for a Quick Convex Hull Algorithm. a) Division in two groups with a line with the minimum and maximum points along x-axis. b) Triangles formed in each group with the points that has the maximum distance to the line. c) Triangles formed with the previous sides and the maximum distance to those sides. d) Final hull found.

Now that we have the hull, we can determine the EER. Since the EER represents the point where FAR and FRR have the same value, we proceed to establish a line with a slope $m = 1$. This is represented by the line $x - y = 0$. The point where this line intersects the hull, is the point where we approximate the EER. We talk about approximation, because we are working with discrete values. Therefore, we do not obtain an exact value, but we do have an approximation. The calculation is the same that we presented in section 7.1.1.

Section 7.2 presents the evaluation of these calculations methods. The previous two sections have presented important tools to implement and validate the Fingerprint and ECG bi-modal authentication algorithm. The following section presents the third contribution of

this work, which is the actual Fingerprint and ECG bi-modal authentication algorithm. The following section would have not been possible to address it correctly without the introduction of the previous to sections that are important for an accurate ECG authentication biometric and an appropriate evaluation method for the proposed bi-modal algorithm.

7.2 Evaluation

This section presents the evaluation of the proposed method to estimate the DET curve with more than one threshold and the mathematical calculation of the EER for non-normal distributions. Section 7.1 previously presented this method and provided the tools to calculate the parameters to evaluate the Fingerprint and ECG bi-modal authentication algorithm. The evaluated methods presented in this chapter, calculates the EER and DET evaluates the bi-modal algorithm of Chapter 5.

The proposed method is evaluated with simulated and real data that represent ECG records as a biometric authentication. First, we evaluate the EER as the intersection of two curves. Then, we use this approach to evaluate the EER in the multi-threshold DET curve.

7.2.1 Experiment Setup

We evaluate our method with four datasets. The first two datasets are the scores that corresponds to 73 subjects from a previous work [11]. These are ECG records from Physionet [85]. Each subject has provided seven samples; one sample is for enrolling and the other six are for authentication. The first set corresponds to scores of genuine users; the total number of genuine scores is 438. This dataset has a mean $\mu = 13.39\%$ and a standard deviation of $\sigma = 8.65\%$. The second dataset corresponds to the scores of imposter users. Similarly, each user has 1 sample for enrolling and 6 samples for authentication; the total number of imposter scores is 31536. The impostor dataset has a mean $\mu = -2.29\%$ and a standard deviation of $\sigma = 3.61\%$. The size of each dataset is the result of authenticating all the users in a batch mode.

We generated the third and fourth datasets with the *randn* function of Matlab. These datasets are similar to the first two datasets in terms of size, mean and standard deviation. The difference is the distribution of data. For these two random generated datasets, we are using a normal distribution (Gaussian). The third data set corresponds to the scores of genuine users (same size, μ and σ as the first dataset) and the fourth dataset corresponds to the scores of imposter users (same size, μ and σ as the second dataset). We generated these last two datasets with normal distribution to calculate the EER with the formula presented by Poh et al. [121]. The obtained results from the formula are our ground truth in order to compare the results of our method.

We evaluate the calculation of the EER explained in section 7.1.1 with the third and fourth datasets. These datasets have normal distribution in order to apply equation 5. We generated datasets 3 and 4 one thousand times in order to run one thousand experiments and have enough information to estimate the errors. More than one thousand experiments do not have a significant difference in the results of the experiments. We used the same mean and standard deviation to generate the data; however, this generated data does not have exactly the same mean and standard deviation that we provide in the settings. The difference is minimal; but in order to have a more accurate estimation we calculate the mean and standard deviation for all the one thousand generated datasets. In each experiment, we provide the calculated mean and standard deviation to equation 5 and calculate the EER in all one thousand experiments. In the next stage we calculate the EER with our method. We use the same datasets to calculate the EER with our intersection of curves method; similarly, we run one thousand experiments and compare both results. We present the results in term of the distribution of the error, as depicted in Figure 28.

To calculate the EER in the multi-threshold approach; first, we obtain the DET curve. We generate two DET curves using two datasets; as before, one real and one generated dataset. The real dataset is the same as [11] and the procedure to generate a dataset with simulated data is the same as the one described at the beginning of this section; where, genuine users has a mean $\mu = 13.39\%$ and a standard deviation $\sigma = 8.65\%$ and imposter users with a mean $\mu = -2.29\%$ and a standard deviation $\sigma = 3.61\%$. The only difference

is that we generate seven thresholds for this part. In the experiment, we calculate the FRR and FAR with independent threshold values. We run the experiments in a batch mode starting from -15% to 35% in steps of 5%.

We present the evaluation of results in two sections, the first one represents the calculation of EER and the second one the DET curve for a multi-threshold approach. In Figure 28 we presented the error distribution of the EER calculation. In Figure 29, we present the DET curve in a multi-threshold approach using simulated data with normal distribution. We show in Figure 30 the DET graph using real data with non-normal distribution.

7.2.2 EER results for intersection of curves

In Figure 28 we can see the calculated error has a normal distribution. The standard deviation is 0.62% and by using the three sigma rule [146] we can say that 68% of the calculated data will have an error in between ±0.62% from the ground truth. 95% of the calculated data will have an error between ±1.24% and 99.7% of the calculated data will have an error between ±1.86%. From a different perspective of the results, we can say that the probability of having an error of ±0.62% is 68%. There is a 27% probability of having an error between 0.62% to 1.24% and -0.62% to -1.24%. There is a 4.7% probability of having an error between 1.24% to 1.86% and -1.24% to -1.86%. Lastly, there is a probability of 0.43% that our method will have and error greater than 1.86%.

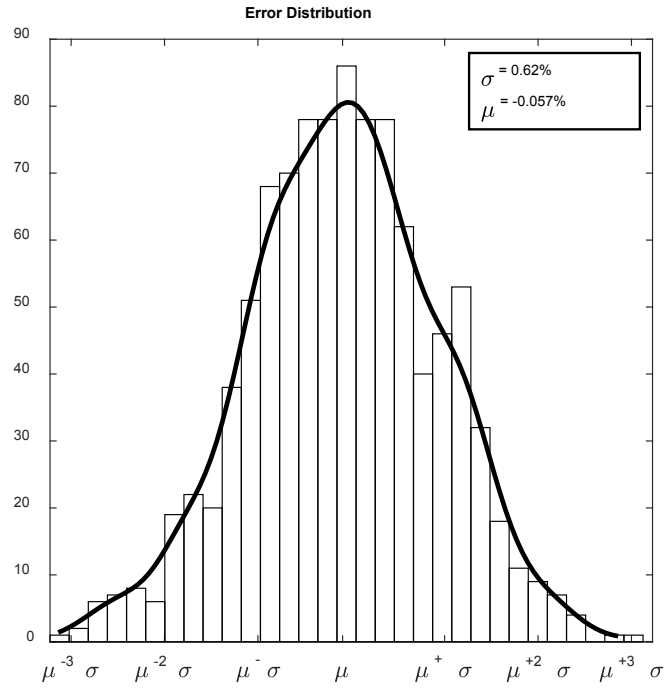

Figure 28. Error distribution of 1000 experiments for the Calculation of EER.

These results proofs that we can use this method in order to calculate the EER. The concept of EER is easy to visualize, this causes that authors obtain the approximated values directly from the DET graph. However, it is important to formalize the methodology to get these results. This method show that the results differ from the eq. (5) in 0.62% at one standard deviation. This difference is not significant considering that this method is applicable to any type of data distribution. In biometrics, this is important, since most of the scores are not normal and will differ depending on the trait and algorithms that can be used.

7.2.3 DET generation and EER results in a multi-threshold biometric

Figure 29 shows the estimated DET with a continuous line and the EER calculated as the intersection of the curves. We can observe that many of the data is redundant and the DET shows the points that should be considered for evaluation of a multi-threshold algorithm.

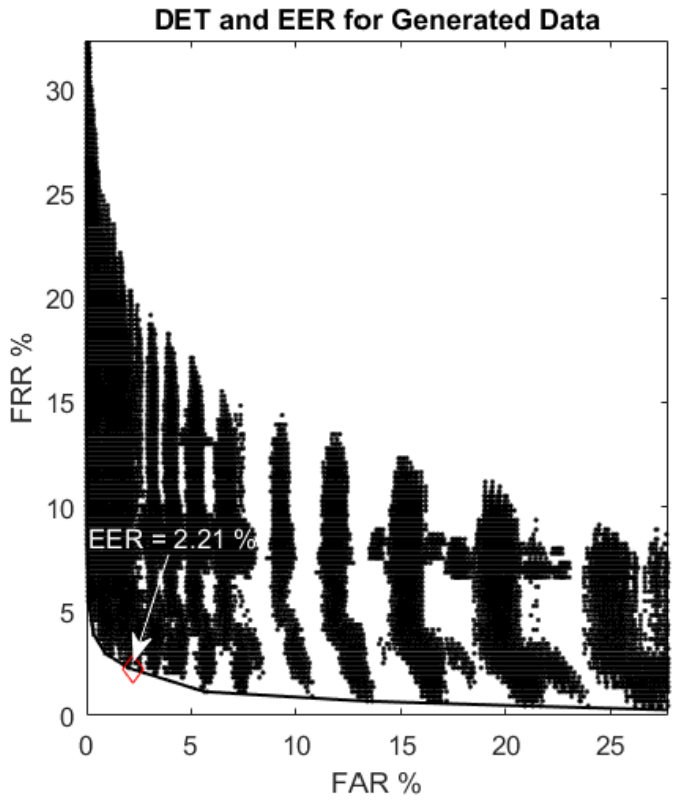

Figure 29. DET and EER in multi-threshold biometric with Generated Data.

Figure 30 shows the results of multi-threshold biometric with real data. We can observe that the behaviour is the same as with the simulated data. The difference in the datasets is the distribution. With the results presented in Figure 29 and Figure 30, we can see that our method can be applied in a normal or non-normal distribution. This last one is the case for biometrics.

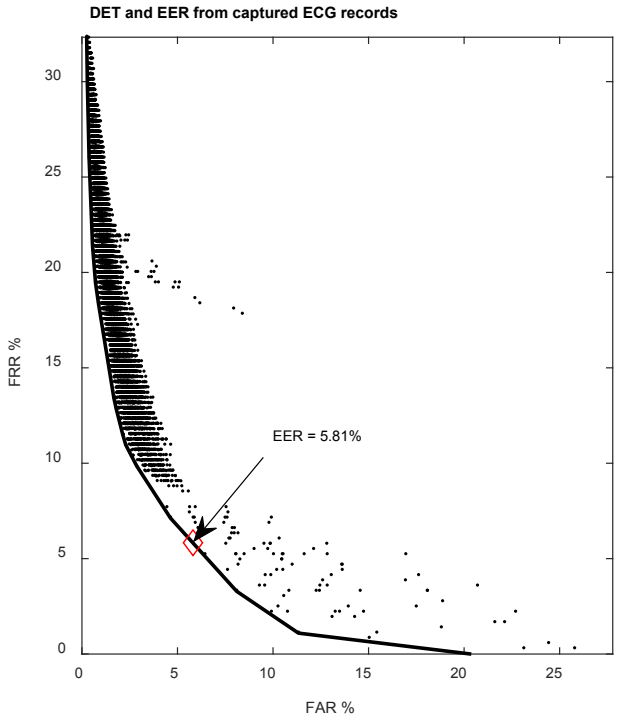

Figure 30. DET and EER in multi-threshold biometric with Real Data.

In biometrics, most of the studies reduce the classification problem to a one threshold variable problem. They do this to reduce the complexity of the algorithms. Nevertheless, having several thresholds helps to increase the performance of the algorithms. The field of multiple thresholds in biometrics has not been widely explored; the literature is limited in terms of calculating EER or plotting the DET curve with multi-threshold. Hence, in this work we present a method to calculate these parameters, to be used when a study requires adjusting more than one threshold in order to obtain a better performance. This is not just limited to multi-threshold; it can also be applied in multibiometrics. Where fusing provides more parameters to be adjust. Results with different settings will lead to a DET with overlapped operating points. However with the proposed method we will be able to find the DET curve and therefore the optimal operating point in the DET curve; for evaluation purposes the optimal operating point is the location of the EER.

Chapter 8.

Conclusions and Future Work

8.1 Conclusions

In this thesis, we addressed the issue of spoofing attacks to biometric systems by duplication of biometric traits. To solve this issue, we designed and developed an ECG biometric authentication system. The advantage of our system is a low acquisition time of 4 seconds and yet, still has accurate results that are comparable with the state of the art works in ECG authentication.

In general, the obtained results of this work show that a deeper understanding of the ECG biometric features leads to better results with shorter acquisition times for authentication. Additionally, fusion of ECG with other biometric modalities improves the accuracy and still maintains a short signal acquisition for authentication. We reached this conclusion based on the different findings of this thesis. The following paragraphs describe the findings of this thesis and present possible directions of future works.

Our first finding is that setting automatic thresholds—on manually established ECG features—improves the accuracy while maintaining 4 seconds of acquisition time for

authentication. To demonstrate that, we designed and developed an algorithm with SVM and obtained a TAR of 100% with a FAR of 7%. These results are an improvement over previous works that use manually tuned thresholds on the ECG features. It is important to mention that reaching a TAR of 100% does not mean that the algorithm is perfect. We have to consider that it has a FAR of 7%. In another words, this algorithm always accept genuine users—never reject them—but will accept 7% of impostors as genuine.

Another finding of this thesis is that automatic detection of ECG features and automatic thresholds on these features, improves accuracy with a short acquisition time of 4 seconds for authentication. This is the result of using CNN—Deep Learning—to automatically extracts features and perform authentication with a one-class classification SVM. We used wavelets transformation to convert an ECG signal into an image and use it with a pre-trained CNN model. One advantage of the pre-trained CNN model is that does not require training. It has been previously trained with millions of data. However, the extracted features from the CNN model needs to train the one-class SVM classifier before performing authentication. With 1 minute of ECG signal to train the SVM model and 4 seconds of ECG for authentication, we obtained an EER of 4.9%. We improved the result to an EER of 2.84% with 2 minutes of ECG for training and maintaining 4 seconds of ECG for authentication. This result represents an improvement over the related work. It is important to mention that none of the related work used 4 seconds of ECG for authentication, but we tested our algorithm with the same procedure they did on the same databases.

Moreover, in this thesis we further researched this topic with a bi-modal fusion approach. This led us to our next finding, where results show that bi-modal fusion of ECG with Fingerprints improves the accuracy with a short acquisition time of ECG authentication. The EER for the bimodal algorithm is 0.46%, which is a better accuracy in contrast to related bi-modal work and the results presented on this thesis. In terms of security, this bi-modal approach uses ECG to reduce the spoofing vulnerability of fingerprint.

R-peak detection is an important process for the ECG authentication. In this thesis, we have designed and developed an R peak detection algorithm with a low average time error. We found that differentiation on the detection of R peaks reduces the average time error. This

characteristic is important in ECG biometrics because a lower average time error, leads to a more precise location of the R-peaks. Thus, this collects ECG features that are more precise and leads to better results. Evaluation shows that this algorithm has similar results in terms sensitivity and predictivity to the related works. However, it has a lower average time error of 4.95 milliseconds.

As part of this research process, we designed a method to determine the DET curve in a bi-modal biometric and a formula to calculate the EER with non-normal distributions. We found that our approach differs in ±0.62% from the formal approach. However, the formal approach is only applicable in normal distributions. Our approach is applicable in normal and non-normal distributions. We needed this approach because biometrics handles non-normal distributed data.

8.2 Future Work

One of the limitations of this thesis is the absence of theoretical research on deep learning. This thesis uses a CNN deep learning model algorithm but does not provide a conceptual explanation of the extracted features. These features improve the accuracy with a short acquisition authentication time but they remain conceptually hidden. In order to further advance this topic, it is important to unveil these features and determine the characteristics of these features; hence, we can extract these features with approaches different to deep learning and provide better solutions for ECG biometric authentication.

The use of better deep learning models can lead to better results. Deep learning is an area of exhaustive research and every day we have new advancements. Pre-trained models with better architectures and more training data will provide better results. This work is still open to try new advancements in deep learning to improve accuracy and reduce the acquisition time for ECG authentication.

In general, all the ECG authentication algorithms can improve their results with a better filtration of the ECG signal and a better R-peak detector. Filtration and R-peak detectors can be studied in individual research works and will provide a great contribution to ECG

authentication. The use of better filtration will remove noise that might cause over fitting. Furthermore, a more accurate R-peak detector will prevent misdetection of heartbeats and provide more information to improve results.

Additionally, further research on multimodal with ECG for authentication can improve results. We have applied fingerprint as a biometric trait and obtained favorable results. As a future work, we can use Face biometrics, Iris Biometrics—among others—and study the combination of modalities. The actual bi-modal results with ECG and fingerprint are highly accurate. We can study the fusion with other modalities to determine improvements in the results and set the limit number of modalities that we can use until the results are not significantly different.

The work on this thesis focused on the authentication algorithm using ECG. However, these algorithms operate with sensitive biometric data. Topics about security and privacy of stored biometric data are very important issues, and it should be properly studied them in a specialized work.

References

[1] A. El Saddik, "Digital Twins: The Convergence of Multimedia Technologies," *IEEE Multimed.*, vol. 25, no. 2, pp. 87–92, Apr. 2018.

[2] A. K. Jain, A. Ross, and S. Pankanti, "Biometrics: A tool for information security," *IEEE Trans. Inf. Forensics Secur.*, vol. 1, no. 2, pp. 125–143, 2006.

[3] G. D. Clark and J. Lindqvist, "Engineering Gesture-Based Authentication Systems," *IEEE Pervasive Comput.*, vol. 14, no. 1, pp. 18–25, 2015.

[4] Ponemon Institute, "Visual Privacy Productivity Study," 2012.

[5] P. J. Phillips, A. Martin, C. L. Wilson, and M. Przybocki, "An introduction evaluating biometric systems," *Computer (Long. Beach. Calif).*, vol. 33, no. 2, pp. 56–63, Feb. 2000.

[6] M. Espinoza, C. Champod, and P. Margot, "Vulnerabilities of fingerprint reader to fake fingerprints attacks," *Forensic Sci. Int.*, vol. 204, no. 1–3, pp. 41–49, 2011.

[7] B. Tan and S. Schuckers, "Spoofing protection for fingerprint scanner by fusing ridge signal and valley noise," *Pattern Recognit.*, vol. 43, no. 8, pp. 2845–2857, 2010.

[8] Y. Yang, L. Wu, G. Yin, L. Li, and H. Zhao, "A Survey on Security and Privacy Issues in Internet-of-Things," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1250–1258, Oct. 2017.

[9] C. Roberts, "Biometric attack vectors and defences," *Comput. Secur.*, vol. 26, no. 1, pp. 14–25, 2007.

[10] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern Recognit.*, vol. 38, no. 1, pp. 133–142, 2005.

[11] J. S. Arteaga-Falconi, H. Al Osman, and A. El Saddik, "ECG Authentication for Mobile Devices," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 3, pp. 591–600, 2016.

[12] D. Maltoni, "Fingerprint Recognition, Overview," in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds. Boston, MA: Springer US, 2015, pp. 664–668.

[13] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: a new approach in human identification," *Instrum. Meas. IEEE Trans.*, vol. 50, no. 3, pp. 808–812, 2001.

[14] G. Wübbeler, M. Stavridis, D. Kreiseler, R.-D. Bousseljot, and C. Elster, "Verification of humans using the electrocardiogram," *Pattern Recognit. Lett.*, vol. 28, no. 10, pp. 1172–1175, 2007.

[15] Y. N. Singh and S. K. Singh, "Evaluation of Electrocardiogram for Biometric Authentication," *J. Inf. Secur.*, vol. 3, no. 1, pp. 39–48, Jan. 2012.

[16] A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee, "Wavelet Distance Measure for Person Identification Using Electrocardiograms," *Instrum. Meas. IEEE Trans.*, vol. 57, no. 2, pp. 248–253, Feb. 2008.

[17] I. Chamatidis, A. Katsika, and G. Spathoulas, "Using deep learning neural networks for ECG based authentication," in *2017 International Carnahan Conference on Security Technology (ICCST)*, 2017, pp. 1–6.

[18] A. Goshvarpour and A. Goshvarpour, “Human identification using information theory-based indices of ECG characteristic points,” *Expert Syst. Appl.*, vol. 127, pp. 25–34, 2019.

[19] A. Goshvarpour and A. Goshvarpour, “Human identification using a new matching Pursuit-based feature set of ECG,” *Comput. Methods Programs Biomed.*, vol. 172, pp. 87–94, 2019.

[20] M. Hammad, S. Zhang, and K. Wang, “A novel two-dimensional ECG feature extraction and classification algorithm based on convolution neural network for human authentication,” *Futur. Gener. Comput. Syst.*, vol. 101, pp. 180–196, 2019.

[21] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, and T. Chen, “Recent advances in convolutional neural networks,” *Pattern Recognit.*, vol. 77, pp. 354–377, 2018.

[22] A. Canziani, A. Paszke, and E. Culurciello, “An Analysis of Deep Neural Network Models for Practical Applications,” *CoRR*, vol. abs/1605.0, 2016.

[23] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going Deeper with Convolutions,” in *Computer Vision and Pattern Recognition (CVPR)*, 2015.

[24] A. Ross, “Multibiometrics,” in *Encyclopedia of Biometrics*, S. Z. Li and A. Jain, Eds. Boston, MA: Springer US, 2009, pp. 967–973.

[25] A. Ross and A. K. Jain, “Biometrics, Overview,” in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds. Boston, MA: Springer US, 2015, pp. 289–294.

[26] A. El Saddik, H. F. Badawi, R. Velazquez, F. Laamarti, R. Gámez Diaz, N. Bagaria, and J. S. Arteaga-Falconi, “Dtwins: A Digital Twins Ecosystem for Health and Well-Being,” *IEEE COMSOC MMTC Commun. - Front.*, vol. 14, no. 2, pp. 39–43, 2019.

[27] A. Ross and A. Jain, “Information fusion in biometrics,” *Pattern Recognit. Lett.*, vol. 24, no. 13, pp. 2115–2125, 2003.

[28] S. A. C. Schuckers, “Liveness Detection: Fingerprint,” in *Encyclopedia of Biometrics*, S. Z. Li and A. Jain, Eds. Boston, MA: Springer US, 2009, pp. 924–931.

[29] F. Agrafioti and D. Hatzinakos, “ECG Based Recognition Using Second Order Statistics,” in *6th Annual Communication Networks and Services Research Conference (cnsr 2008)*, 2008, pp. 82–87.

[30] C. C. Chiu, C. M. Chuang, and C. Y. Hsu, “A Novel Personal Identity Verification Approach Using a Discrete Wavelet Transform of the ECG Signal,” in *2008 International Conference on Multimedia and Ubiquitous Engineering (mue 2008)*, 2008, pp. 201–206.

[31] G. G. Molina, F. Bruekers, C. Presura, M. Damstra, and M. van der Veen, “Morphological synthesis of ECG signals for person authentication,” in *2007 15th European Signal Processing Conference*, 2007, pp. 738–742.

[32] A. Lourenço, H. Silva, and A. Fred, “ECG-based biometrics: A real time classification approach,” in *2012 IEEE International Workshop on Machine Learning for Signal Processing*, 2012, pp. 1–6.

[33] Y. N. Singh and S. K. A. Singh, “Identifying Individuals Using Eigenbeat Features of Electrocardiogram,” *J. Eng.*, vol. 2013, no. 539284, p. 8, 2013.

[34] A. Ross and A. K. Jain, “Biometrics, Overview,” in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds. Boston, MA: Springer US, 2015, pp. 289–294.

[35] A. K. Jain, A. Ross, and S. Prabhakar, “An introduction to biometric recognition,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 1, pp. 4–20, Jan. 2004.

[36] J. L. Wayman, “Biometric Verification/Identification/Authentication/Recognition: The Terminology,” in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds. Boston, MA: Springer US, 2015, pp. 263–268.

[37] National Research Council, *Who Goes There?: Authentication Through the Lens of Privacy*. Washington, DC: The National Academies Press, 2003.

[38] S. K. Dahel and Q. Xiao, “Accuracy performance analysis of multimodal biometrics,” in *Information Assurance Workshop, 2003. IEEE Systems, Man and Cybernetics Society*, 2003, pp. 170–173.

[39] J. Arteaga-Falconi, D. P. Tobón, and A. E. Saddik, “EER Calculation and DET Approximation in a Multi-Threshold Biometric System,” in *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2018, pp. 1–6.

[40] S. M. Prasanna, S. Sahoo, and T. Choubisa, “Multimodal Biometric Person Authentication : A Review,” in *IETE Technical Review*, 2012, vol. 29, no. 1, pp. 54–75.

[41] J. S. Arteaga-Falconi, H. Al Osman, and A. El Saddik, “ECG and Fingerprint Bimodal Authentication,” *Sustain. Cities Soc.*, vol. 40, pp. 274–283, 2018.

[42] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki, “The DET curve in assessment of detection task performance,” in *Proc. Eurospeech ’97*, 1997, pp. 1895–1898.

[43] D. Berrar and P. Flach, “Caveats and pitfalls of ROC analysis in clinical microarray research (and how to avoid them),” *Brief. Bioinform.*, vol. 13, no. 1, pp. 83–97, 2012.

[44] R. Cappelli, D. Maio, D. Maltoni, and J. L. Wayman, “Performance evaluation of fingerprint verification systems,” *Pattern Anal. Mach. Intell. IEEE Trans.*, vol. 28, no. 1, pp. 3–18, 2006.

[45] R. V Yampolskiy and V. Govindaraju, “Behavioural biometrics: a survey and classification,” *Int. J. Biom.*, vol. 1, no. 1, pp. 81–113, 2008.

[46] M. O. Derawi, C. Nickel, P. Bours, and C. Busch, “Unobtrusive User-Authentication on Mobile Phones Using Biometric Gait Recognition,” in *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on*, 2010, pp. 306–311.

[47] F. Monrose and A. D. Rubin, “Authentication via keystroke dynamics,” *Proc. 4th ACM Conf. Comput. Commun. Secur. - CCS ’97*, pp. 48–56, 1997.

[48] L. L. Lee, T. Berger, and E. Aviczer, “Reliable on-line human signature verification systems,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 6, pp. 643–647, 1996.

[49] M. Fouad, F. Alsulaiman, A. El Saddik, and E. Petriu, “Revocable handwritten signatures with haptic information,” *HAVE 2011 - IEEE Int. Symp. Haptic Audio-v. Environ. Games, Proc.*, pp. 108–111, 2011.

[50] V. S. Nalwa, “Automatic on-line signature verification,” *Proc. IEEE*, vol. 85, no. 2, pp. 215–239, 1997.

[51] A. El Saddik, M. Orozco, Y. Asfaw, S. Shirmohammadi, and A. Adler, “A Novel Biometric System for Identification and Verification of Haptic Users,” *IEEE Trans. Instrum. Meas.*, vol. 56, no. 3, pp. 895–906,

Jun. 2007.

[52] A. Ross and A. K. Jain, “Human recognition using biometrics: an overview,” *Ann. Des Télécommunications*, vol. 62, no. 1–2, pp. 11–35, 2007.

[53] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, “Comparison and combination of ear and face images in appearance-based biometrics,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1160–1165, 2003.

[54] D. Zhang, W. K. Kong, J. You, and M. Wong, “Online palmprint identification,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1041–1050, 2003.

[55] R. Zunkel, “Hand geometry based authentication,” *Biometrics Pers. Identif. Networked Soc.*, pp. 87–102, 1999.

[56] D. Ho Cho, K. R. Park, D. W. Rhee, Y. Kim, and J. Yang, “Pupil and Iris Localization for Iris Recognition in Mobile Phones,” in *Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2006. SNPD 2006. Seventh ACIS International Conference on*, 2006, pp. 197–201.

[57] F. Scotti and V. Piuri, “Adaptive reflection detection and location in iris biometric images by using computational intelligence techniques,” *Instrum. Meas. IEEE Trans.*, vol. 59, no. 7, pp. 1825–1833, 2010.

[58] J. A. Markowitz, “Voice Biometrics,” *Commun. ACM*, vol. 43, no. 9, pp. 66–73, Sep. 2000.

[59] S. Wang and J. Liu, “Biometrics on Mobile Phone,” in *Recent Applications in Biometrics*, D. J. Yang, Ed. InTech, 2011, pp. 3–22.

[60] J. Bigun, “Fingerprint Features,” in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds. Boston, MA: Springer US, 2015, pp. 609–619.

[61] I. Iancu and N. Constantinescu, “Intuitionistic fuzzy system for fingerprints authentication,” *Appl. Soft Comput.*, vol. 13, no. 4, pp. 2136–2142, 2013.

[62] K. Cao, X. Yang, X. Chen, Y. Zang, J. Liang, and J. Tian, “A novel ant colony optimization algorithm for large-distorted fingerprint matching,” *Pattern Recognit.*, vol. 45, no. 1, pp. 151–161, 2012.

[63] X. Tan and B. Bhanu, “Fingerprint matching by genetic algorithms,” *Pattern Recognit.*, vol. 39, no. 3, pp. 465–477, 2006.

[64] Z. M. Kovacs-Vajna, “A fingerprint verification system based on triangular matching and dynamic time warping,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1266–1276, Nov. 2000.

[65] C. I. Watson, M. D. Garris, E. Tabassi, C. L. Wilson, R. M. McCabe, S. Janet, and K. Ko, “User’s Guide to NIST Biometric Image Software (NBIS),” *National Institute of Standards and Technology*, no. November. National Institute of Standards and Technology, 2008.

[66] S. Maddala, S. R. Tangellapally, J. S. Bartůněk, and M. Nilsson, “Implementation and evaluation of NIST Biometric Image Software for fingerprint recognition,” in *ISSNIP Biosignals and Biorobotics Conference 2011*, 2011, pp. 1–5.

[67] A. L. Goldberger, *Clinical electrocardiography: a simplified approach*, 8th ed. Mosby Elsevier, 2012.

[68] R. Hoekema, G. J. H. Uijen, and A. Van Oosterom, “Geometrical aspects of the interindividual variability of multilead ECG recordings,” *Biomed. Eng. IEEE Trans.*, vol. 48, no. 5, pp. 551–559, May 2001.

[69] A. Van Oosterom, R. Hoekema, and G. J. Uijen, “Geometrical factors affecting the interindividual

variability of the ECG and the VCG.," *J. Electrocardiol.*, vol. 33, pp. 219–227, 2000.

[70] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ECG: Convolutional Neural Networks for ECG biometric recognition," *Pattern Recognit. Lett.*, 2018.

[71] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet.," *Biomed. Tech. Biomed. Eng.*, vol. 40, no. 1, pp. 317–318, 1995.

[72] E. J. da Silva Luz, G. J. P. Moreira, L. S. Oliveira, W. R. Schwartz, and D. Menotti, "Learning Deep Off-the-Person Heart Biometrics Representations," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 5, pp. 1258–1270, May 2018.

[73] Y. Chu, H. Shen, and K. Huang, "ECG Authentication Method Based on Parallel Multi-Scale One-Dimensional Residual Network With Center and Margin Loss," *IEEE Access*, vol. 7, pp. 51598–51607, 2019.

[74] Z. Zhao, Y. Zhang, Y. Deng, and X. Zhang, "ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation," *Comput. Biol. Med.*, vol. 102, pp. 168–179, 2018.

[75] H. B. Mitchell, "Multi-sensor data fusion: An introduction," *Multi-Sensor Data Fusion An Introd.*, pp. 1–281, 2007.

[76] E. J. C. Kelkboom, X. Zhou, J. Breebaart, R. N. J. Veldhuis, and C. Busch, "Multi-algorithm fusion with template protection," *IEEE 3rd Int. Conf. Biometrics Theory, Appl. Syst. BTAS 2009*, 2009.

[77] B. Ulery, R. A. Hicklin, C. Watson, W. Fellner, and P. Hallinan, "Studies of Biometric Fusion," *NIST Tech. Rep. IR7346*, 2006.

[78] K. I. Chang, K. W. Bowyer, and P. J. Flynn, "An evaluation of multimodal 2d+3d face biometrics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 4, pp. 619–624, 2005.

[79] S. C. Dass, K. Nandakumar, and A. K. Jain, "A Principled Approach to Score Level Fusion in Multimodal Biometric Systems," *Audio- Video-Based Biometric Pers. Authentication, Lect. Notes Comput. Sci.*, no. i, pp. 1049–1058, 2005.

[80] S. Prabhakar and A. K. Jain, "Decision-level fusion in fingerprint verification," *Pattern Recognit.*, vol. 35, no. 4, pp. 861–874, 2002.

[81] P. Mordohai, G. Medioni, P. Fua, A. Ross, J. Soh, F. Deravi, A. Triglia, A. Bazin, F. Roli, J. Bigun, B. Roui-Abidi, and M. Abidi, "Multiple Classifiers," in *Encyclopedia of Biometrics*, S. Z. Li and A. Jain, Eds. Boston, MA: Springer US, 2009, pp. 986–986.

[82] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli, "Multimodal fusion for multimedia analysis: a survey," *Multimed. Syst.*, vol. 16, no. 6, pp. 345–379, Nov. 2010.

[83] Š. Raudys and F. Roli, "The Behavior Knowledge Space Fusion Method: Analysis of Generalization Error and Strategies for Performance Improvement," in *Multiple Classifier Systems: 4th International Workshop, MCS 2003 Guildford, UK, June 11--13, 2003 Proceedings*, T. Windeatt and F. Roli, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 55–64.

[84] E. Manjunathswamy, A. M. Abhishek, J. Thriveni, R. Venugopal, and L. Patnaik, "Multimodal Biometric

Authentication using ECG and Fingerprint," *Int. J. Comput. Appl.*, vol. 111, no. 13, pp. 33–39, Feb. 2015.

[85] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, p. e215 LP-e220, Jun. 2000.

[86] National Institute of Standard and Technology, "Biometric Score Set," 2011. [Online]. Available: http://www.itl.nist.gov/iad/894.03/biometricscores/.

[87] D. Griffin and Jae Lim, "Signal estimation from modified short-time Fourier transform," *IEEE Trans. Acoust.*, vol. 32, no. 2, pp. 236–243, Apr. 1984.

[88] M. K. Kıymık, İ. Güler, A. Dizibüyük, and M. Akın, "Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application," *Comput. Biol. Med.*, vol. 35, no. 7, pp. 603–616, 2005.

[89] C. Chakrabarti, M. Vishwanath, and R. M. Owens, "Architectures for wavelet transforms: A survey," *J. VLSI signal Process. Syst. signal, image video Technol.*, vol. 14, no. 2, pp. 171–192, Nov. 1996.

[90] J. Allen, "Short term spectral analysis, synthesis, and modification by discrete Fourier transform," *IEEE Trans. Acoust.*, vol. 25, no. 3, pp. 235–238, Jun. 1977.

[91] B. Jawerth and W. Sweldens, "An Overview of Wavelet Based Multiresolution Analyses," *SIAM Rev.*, vol. 36, no. 3, pp. 377–412, 1994.

[92] J. M. Lilly and S. C. Olhede, "Generalized Morse Wavelets as a Superfamily of Analytic Wavelets," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 6036–6041, Nov. 2012.

[93] J. Lin and L. Qu, "Feature extraction based on morlet wavelet and its application for mechanical fault diagnosis," *J. Sound Vib.*, vol. 234, no. 1, pp. 135–148, 2000.

[94] X. Mi, H. Ren, Z. Ouyang, W. Wei, and K. Ma, "The use of the Mexican Hat and the Morlet wavelets for detection of ecological patterns," *Plant Ecol.*, vol. 179, no. 1, pp. 1–19, Jul. 2005.

[95] M. Holschneider and I. Iglewska-Nowak, "Poisson Wavelets on the Sphere," *J. Fourier Anal. Appl.*, vol. 13, no. 4, pp. 405–419, Aug. 2007.

[96] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, p. 436, May 2015.

[97] Y. Kim, "Convolutional Neural Networks for Sentence Classification," *arXiv e-prints*, p. arXiv:1408.5882, Aug. 2014.

[98] I. Krasin, T. Duerig, N. Alldrin, V. Ferrari, S. Abu-El-Haija, A. Kuznetsova, H. Rom, J. Uijlings, S. Popov, A. Veit, S. Belongie, V. Gomes, A. Gupta, C. Sun, G. Chechik, D. Cai, Z. Feng, D. Narayanan, and K. Murphy, "OpenImages: A public dataset for large-scale multi-label and multi-class image classification.," *Dataset available from https//github.com/openimages*, 2017.

[99] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1717–1724.

[100] J. Schmidt-Hieber, "Nonparametric regression using deep neural networks with ReLU activation function," *arXiv e-prints*, p. arXiv:1708.06633, Aug. 2017.

[101] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep Recurrent Neural Networks for Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639–3655, Jul. 2017.

[102] I. Sutskever and G. E. Hinton, "Deep, narrow sigmoid belief networks are universal approximators," *Neural Comput.*, vol. 20, no. 11, pp. 2629–2636, 2008.

[103] H. Wu and X. Gu, "Towards dropout training for convolutional neural networks," *Neural Networks*, vol. 71, pp. 1–10, 2015.

[104] V. Suárez-Paniagua and I. Segura-Bedmar, "Evaluation of pooling operations in convolutional architectures for drug-drug interaction extraction," *BMC Bioinformatics*, vol. 19, no. 8, p. 209, 2018.

[105] D. Mishkin, N. Sergievskiy, and J. Matas, "Systematic evaluation of convolution neural network advances on the Imagenet," *Comput. Vis. Image Underst.*, vol. 161, pp. 11–19, 2017.

[106] M. Lin, Q. Chen, and S. Yan, "Network In Network," *arXiv e-prints*, p. arXiv:1312.4400, Dec. 2013.

[107] F. Lauer, C. Y. Suen, and G. Bloch, "A trainable feature extractor for handwritten digit recognition," *Pattern Recognit.*, vol. 40, no. 6, pp. 1816–1824, 2007.

[108] W. Liu, Y. Wen, Z. Yu, and M. Yang, "Large-Margin Softmax Loss for Convolutional Neural Networks," *arXiv e-prints*, p. arXiv:1612.02295, Dec. 2016.

[109] N. M. Arzeno, Z.-D. Deng, and C.-S. Poon, "Analysis of First-Derivative Based QRS Detection Algorithms," *Biomed. Eng. IEEE Trans.*, vol. 55, no. 2, pp. 478–484, Feb. 2008.

[110] D. Benitez, P. A. Gaydecki, A. Zaidi, and A. P. Fitzpatrick, "The use of the Hilbert transform in ECG signal analysis," *Comput. Biol. Med.*, vol. 31, no. 5, pp. 399–406, 2001.

[111] P. S. Hamilton and W. J. Tompkins, "Quantitative Investigation of QRS Detection Rules Using the MIT/BIH Arrhythmia Database," *IEEE Trans. Biomed. Eng.*, vol. BME-33, no. 12, pp. 1157–1165, 1986.

[112] B.-U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," *Eng. Med. Biol. Mag. IEEE*, vol. 21, no. 1, pp. 42–57, Jan. 2002.

[113] J. Jang and H. Kim, "Performance Measures," in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds. Boston, MA: Springer US, 2015, pp. 1230–1237.

[114] M. He, S.-J. Horng, P. Fan, R.-S. Run, R.-J. Chen, J.-L. Lai, M. K. Khan, and K. O. Sentosa, "Performance evaluation of score level fusion in multimodal biometric systems," *Pattern Recognit.*, vol. 43, no. 5, pp. 1789–1800, 2010.

[115] C.-H. Chen and C. Te Chu, "Fusion of Face and Iris Features for Multimodal Biometrics," in *Advances in Biometrics: International Conference, ICB 2006, Hong Kong, China, January 5-7, 2006. Proceedings*, D. Zhang and A. K. Jain, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 571–580.

[116] R. Brunelli and D. Falavigna, "Person identification using multiple cues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 10, pp. 955–966, 1995.

[117] L. Hong and A. Jain, "Integrating faces and fingerprints for personal identification," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 1351, pp. 16–23, 1997.

[118] U. Dieckmann, P. Plankensteiner, and T. Wagner, "SESAM: A biometric person identification system using sensor fusion," *Pattern Recognit. Lett.*, vol. 18, no. 9, pp. 827–833, 1997.

[119] A. Nagar, K. Nandakumar, and A. K. Jain, “Multibiometric cryptosystems based on feature-level fusion,” *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 1 PART 2, pp. 255–268, 2012.

[120] Y. Kim, K. A. Toh, A. B. J. Teoh, H. L. Eng, and W. Y. Yau, “An online learning network for biometric scores fusion,” *Neurocomputing*, vol. 102, pp. 65–77, 2013.

[121] N. Poh and S. Bengio, “Why do multi-stream, multi-band and multi-modal approaches work on biometric user authentication tasks?,” *2004 IEEE Int. Conf. Acoust. Speech, Signal Process.*, vol. 5, pp. 893–896, 2004.

[122] A. K. Jain and A. Ross, “Introduction to Biometrics,” in *Handbook of Biometrics*, A. K. Jain, P. Flynn, and A. A. Ross, Eds. Boston, MA: Springer US, 2008, pp. 1–22.

[123] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “PhysioBank, PhysioToolkit, and PhysioNet : Components of a New Research Resource for Complex Physiologic Signals,” *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[124] G. B. Moody and R. G. Mark, “The impact of the MIT-BIH Arrhythmia Database,” *Eng. Med. Biol. Mag. IEEE*, vol. 20, no. 3, pp. 45–50, May 2001.

[125] A. Taddei, G. Distante, M. Emdin, P. Pisani, G. B. Moody, C. Zeelenberg, and C. Marchesi, “The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography,” *Eur. Heart J.*, vol. 13, no. 9, pp. 1164–1172, 1992.

[126] P. Laguna, R. G. Mark, A. Goldberg, and G. B. Moody, “A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG,” in *Computers in Cardiology 1997*, 1997, pp. 673–676.

[127] J. A. van Alsté, W. van Eck, and O. E. Herrmann, “ECG baseline wander reduction using linear phase filters,” *Comput. Biomed. Res.*, vol. 19, no. 5, pp. 417–427, 1986.

[128] S. Orfanidis, *Introduction to Signal Processing*. Rutgers University, 2010.

[129] J. S. Arteaga-Falconi, H. Al Osman, and A. El Saddik, “R-peak detection algorithm based on differentiation,” in *Intelligent Signal Processing (WISP), 2015 IEEE 9th International Symposium on*, 2015, pp. 1–4.

[130] J. Ribeiro Pinto, J. S. Cardoso, and A. Lourenço, “Evolution, Current Challenges, and Future Possibilities in ECG Biometrics,” *IEEE Access*, vol. 6, pp. 34746–34776, 2018.

[131] J. M. Lilly and S. C. Olhede, “Higher-Order Properties of Analytic Wavelets,” *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 146–160, Jan. 2009.

[132] D. Borland and R. M. Taylor Ii, “Rainbow Color Map (Still) Considered Harmful,” *IEEE Comput. Graph. Appl.*, vol. 27, no. 2, pp. 14–17, Mar. 2007.

[133] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[134] S. S. Khan and M. G. Madden, “A Survey of Recent Trends in One Class Classification,” in *Artificial Intelligence and Cognitive Science*, 2010, pp. 188–197.

[135] D. Tax, “One-class classification,” Delft University of Technology, 2001.

[136] Yunqiang Chen, Xiang Sean Zhou, and T. S. Huang, “One-class SVM for learning in image retrieval,” in *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*, 2001, vol. 1, pp. 34–37 vol.1.

[137] B. I. D. M. Center, “MIT-BIH Normal Sinus Rhythm Database.” [Online]. Available: http://physionet.org/physiobank/database/nsrdb/.

[138] H. P. da Silva, A. Lourenço, A. Fred, N. Raposo, and M. Aires-de-Sousa, “Check Your Biosignals Here: A new dataset for off-the-person ECG biometrics,” *Comput. Methods Programs Biomed.*, vol. 113, no. 2, pp. 503–514, 2014.

[139] R. Cappelli, M. Ferrara, A. Franco, and D. Maltoni, “Fingerprint verification competition 2006,” *Biometric Technol. Today*, vol. 15, no. 7–8, pp. 7–9, 2007.

[140] A. K. Jain, A. A. Ross, and K. Nandakumar, “Introduction,” in *Introduction to Biometrics*, Boston, MA: Springer US, 2011, pp. 1–49.

[141] R. A. Robergs and R. Landwehr, “The surprising history of the ‘HRmax=220-age’ equation,” *J. Exerc. Physiol. Online*, vol. 5, no. 2, pp. 1–10, 2002.

[142] I. Silva and G. Moody, “An Open-source Toolbox for Analysing and Processing PhysioNet Databases in MATLAB and Octave,” *J. Open Res. Softw.*, vol. 2, no. 1, 2014.

[143] A. R. P. N. Standard, *ANSI/AAMI EC57: Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms*. 1998.

[144] “Line Segment Intersection,” in *Computational Geometry: Algorithms and Applications*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 19–43.

[145] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, “The Quickhull Algorithm for Convex Hulls,” *ACM Trans. Math. Softw.*, vol. 22, no. 4, pp. 469–483, Dec. 1996.

[146] E. Grafarend, “A first confidence interval of Gauss-Laplace normally distributed observations: μ and σ^2 known the three sigma rule,” in *Linear and nonlinear models: fixed effects, random effects, and mixed models*, Stuttgart, Germany, 2006, p. 773.