FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Don't You Forget About Me: Enhancing Long-Term Performance in Electrocardiogram Biometrics

Gabriel Lopes

Mestrado Integrado em Bioengenharia

Supervisor: Jaime dos Santos Cardoso, PhD

Co-Supervisor: João Ribeiro Pinto, MSc

July 17, 2019

Don't You Forget About Me: Enhancing Long-Term Performance in Electrocardiogram Biometrics

Gabriel Lopes

Mestrado Integrado em Bioengenharia

Approved in public examination by the Jury:

President: Miguel Velhote Correia, PhD

Opponent: Susana Brás, PhD

Referee: Jaime S. Cardoso, PhD

July 17, 2019

Resumo

Hoje em dia, é esperado assegurar a integridade física, mental ou moral de um individuo. Num mundo coberto de tecnologia, a violação desta integridade é facilitada, pelo que ataques de contrafação/ roubo de identidade são comuns. Embora os métodos atuais de segurança visem dificultar estes ataques, ainda não existe um sistema infálivel. Um dos método é a biometria (técnica de reconhecimento do indivíduo) que contém algumas vantagens relativamente aos metodos tradicionais tais como ser conveniente, redução de custos e aumento da segurança.

Esta dissertação focou na implementação de processos/ técnicas do estado-da-arte para melhorar a performance de um sistema biométrico de eletrocardiograma em condições realistas durante um longo periodo de tempo. Distintos desafios foram encontrados para procurar a melhor solução. Foram utilizadas diversas abordagens no pré-processamento bem como na extração de caracteristicas para posterior avaliação.

Recorrendo a um filtro passa banda (1-40 Hz), com a aplicação da DCT na média dos complexos QRS (janela de 0.35 s) e com o kNN como classificador obteram-se as melhores melhorias em média com a técnica de fixação (j $\times$ 3 + 3) Os resultados revelaram uma melhoria de 10.0%, que apesar de ser uma melhoria significativa comparativamente com os metodos do estado-da-arte, ainda não é o suficiente para aplicações de um longo periodo de tempo.

Abstract

Nowadays, the physical, mental or moral integrity of an individual is expected to be ensured. In a technology-covered world, violation of this integrity is facilitated, so counterfeit/identity theft attacks are common. Although the current methods of security, that aim to hamper this attack, there is still no infallible system. One method is biometrics (individual recognition technique) which contains some advantages over traditional methods such as convenience, reduced costs, and increased safety technique.

This dissertation focused on the implementation of state-of-the-art processes/techniques to improve the performance of a biometric electrocardiogram system under realistic conditions over a long period of time. Different challenges were found to find the best solution. Several approaches were used in the pre-processing as well as in the extraction of characteristics for later evaluation.

Using a bandpass filter (1-40 Hz), with the application of the DCT in the mean of the QRS complexes (window of 0.35 s) and with the kNN as the classifier, the best improvements were obtained on average with the fixation technique (j $\times$ 3 + 3) The results showed an improvement of 10.0%, which despite being a significant improvement compared to the state-of-the-art methods, is still not enough for applications over a long period of time.

Acknowledgements

Contents

List of Figures

List of Tables

Abbreviations

2D	Bi dimensional
AC	Alternating Current
ACC	Accuracy
ACL	Autocorrelation
ANN	Artificial Neural Network
ANOVA	One-way Analysis of Variance
AV	Atrioventricular
aVF	Augmented Vector Foot
aVR	Augmented Vector Right Wrist
aVL	Augmented Vector Left Wrist
AUC	Area Under the Curve
BPF	Bandpass Filter
BVP	Blood Volume Pressure
CFS	Correlation feature selection
CMC	Cumulative Match Characteristic
cNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
DCT	Discrete Cosine Transform
DET	Decision Error Tradeoff
DNA	Deoxyribonucleic Acid
ECG	Electrocardiogram
EER	Equal Error Rate
EFDT	Extremely Fast Decision Tree
FAR	False Acceptance Rate
FLDA	Fisher Linear Discriminant Analysis
FN	False Negative
FNIR	False-Negative Identification Rate
FP	False Positive
FPF	False Positive Fraction
FPIR	False Positive Identification Rate
FRR	False Reject Rate
GBFS	Greedy Best First Search
GUMR	Genuine Update Miss Rate
HLDA	Heteroscedastic Linear discriminant Analysis
HMM	Hidden Markov Model
KPCA	Kernel Principal Component Analysis
HRV	Heart Rate Variability
IDR	Accuracy

IUSR	Impostor Update Selection Rate
kNN	k-Nearest Neighbours
LBP	Local Binary Patterns
LDA	Linear Discriminant Analysis
LLR	Log-Likelihood Ratio
LT	Long-Term
MIDR	Misidentification Rate
ML	Machine Learning
MLP	Multilayer Perceptron
MRE	Mean Error
PCA	Principal Component Analysis
PCG	Phonocardiogram
PIN	Personal Identification Number
PPG	Photoplethysmogram
PTCR	Probability of Time to Correct Reject
RFECV	Recursive feature selection cross validation
ROC	Receiver Operating Characteristic
SA	Sino-Atrial
SFS	Sequential forward selection
SHLDA	Smoothed Heteroscedastic Linear discriminant Analysis
ST	Short-Term
SVM	Support Vector Machine
TCR	Time to Correct Reject
TN	True Negative
TP	True Positive
TPF	True Positive Fraction
TPIR	True Positive Identification Rate
USC	Usability-Security Characteristic Curve
VCG	Vectorcardiography

Chapter 1

Introduction

As a human, one critical thing in life is the identification/recognition of another human being. With the increase of population, society felt the need of establishing a secure system, like a unique ID. A lot of methods are used nowadays such as PIN code, password, ID-card, keys (Prabhakar et al., 2003).

With some disagreement between societies for various reason (as an example, religious) people tend to hurt each other, as such, these systems are susceptible to be copied or counterfeit (Hadid et al., 2015). Also, it is very easy for the person to forget, share or lose them. Stronger institutions like Google, Apple, banks, airport security, and military industry, felt the need to protect more their data. They use something that can characterize the person and belong only to that person - a biometric trait (Agrafioti et al., 2012). The biometric traits, for example, fingerprint, voice, face recognition, electrocardiogram requires that the individual is present when they want to validate because they characterize the human anatomically and the physiological behaviour (Akhtar et al., 2015).

When comparing biometric systems to the traditional techniques, biometrics bring some advantages in identification and authentication of the person, because these are difficult to counterfeit or steal, and are easier to use because all the person needs are their body (Jain et al., 2004).

Through the existing way of acquiring the electrocardiogram (ECG) signal makes this biometric system needs effortless for the user. In this work, we will focus on long-term electrocardiogram, where, according to literature good results with medical signals are obtained, but the performance decay over time (Komeili et al., 2018) (Ye et al., 2010). Hereupon, a big gap exists for improving these systems, with the possibility to put a reliable product in the market. To improve these measures and allow future comparison between the several methods, several template update techniques will be tested.

This dissertation focuses on long-term ECG biometrics and it includes the characterization of the electrocardiography signal (see Chapter 2) and the fundamental concepts of machine learning (see Chapter 3). Chapter 4 contains the biometric systems as well as their differences, Chapter 5 addresses the most common long-term ECG databases and the state-of-the-art regarding template update in biometric systems with ECG. Chapter 6 present a study on the variability and permanence of ECG signal, and exhibit the preprocessing, preparation, feature extraction, and recognition methods. In Chapter 7 the experiments and the results of template update techniques are expose. Finally Chapter 8, adds some conclusions related to the work performed, and discusses some line-thoughts to be explored in the future.

Chapter 2

Electrocardiographic Signal Characterization

2.1 Anatomy and Physiology

The heart assures blood circulation in the body (see Fig. 2.1). We consider the heart a symbol of life and call it the life pump/motor. Depending on gender, a healthy human heart normally weights between 230 to 340 grams. It is found, between the lungs, in the mediastinum (middle compartment of the chest) (Tortora and Derrickson, 2016).

Since humans are mammals, they have a double circuit separated by a septum. One of the circuits is pulmonary circulation, in which the oxygen-poor blood is sent to the lungs where it is oxygenated. To achieve that, flowing through vena cava, the blood arrives at the heart right atrium, is directed to the right ventricle and there is pumped to the lung through the pulmonary arteries.

When it reaches the lungs, the oxygen replaces the carbon dioxide in the red blood cells (pulmonary hematosis). That is allowed because the hemoglobin, have the hemegroup with iron that binds temporarily to oxygen and enables them to transport the oxygen from the lungs throughout the body (Jameson, 2018). The red blood cells enter the heart through the pulmonary veins to the left atrium and follow the path to the left ventricle, where it is pumped to the rest of the body, through the aorta. Myocardium stays between the ventricles. It is more developed in the left ventricle because needs to pump the blood to the body periphery. A double layer sac called pericardium encases, protect and fix the heart inside the thoracic cavity. The heart is lubricated during contractions/movements of the lungs and diaphragm by the pericardial fluid that stays between the outer layer (parietal pericardium), and the inner layer (serous pericardium) (Malasri and Wang, 2009).

A triple-walled constitute the heart's external wall. Epicardium (the outermost wall layer), belongs to the inner wall of the pericardium. Myocardium (the middle layer) consists of the mass of cells that contracts. Endocardium (the inner layer) is the line of cells that interact directly with the blood (Tortora and Derrickson, 2016).

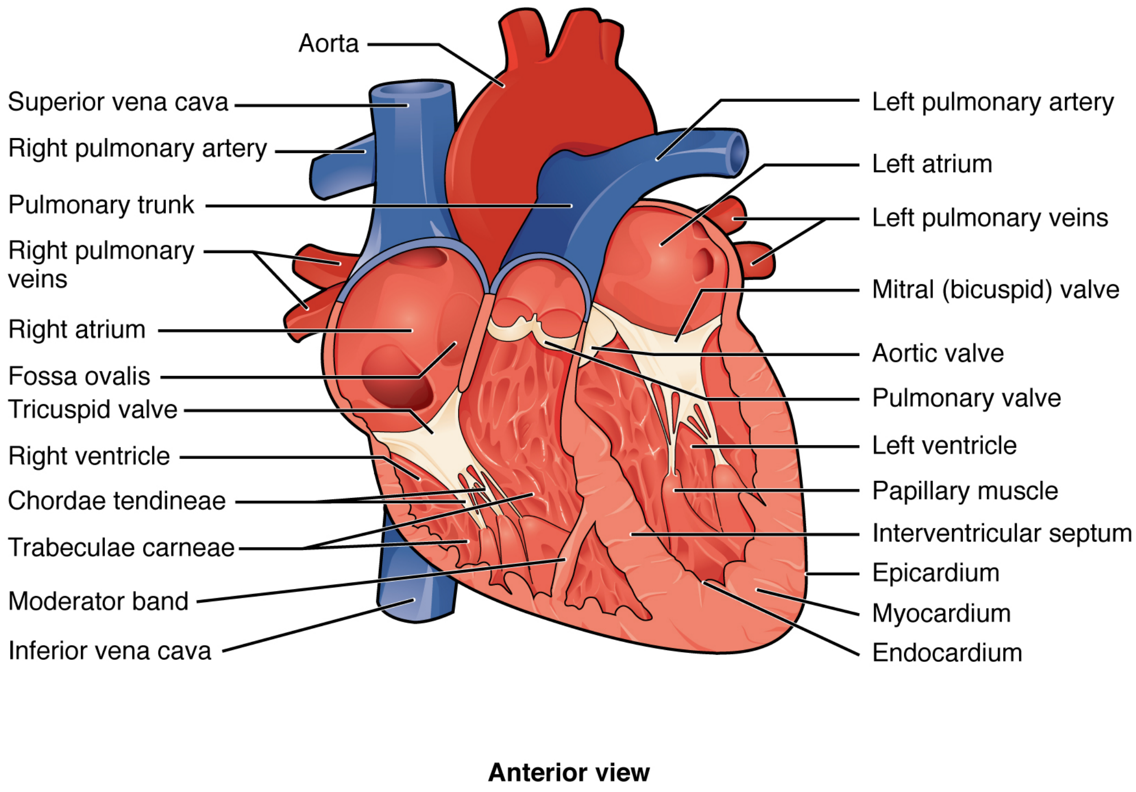


Figure 2.1: Heart description (Betts et al., 2017)

The atrioventricular (AV) valves are constituted by the tricuspid and mitral valves, that made the connection between the atria and the ventricles. The pulmonary artery stays separated from the right ventricle through the pulmonary semi-lunar valve and the aorta stays separated from the left ventricle through the aortic valve. The valves are anchored to heart muscles athwart *chordae tendinea*, or the heartstrings (Seeley et al., 1991).

The heart's activity is coordinated by electrical impulses which allows the perfect synchronization of the muscles in order to pump blood throughout the body at the right rate (time, direction and pressure).

Above the right atrium stays the heart's pacemaker, the Sino-atrial (or sinus, SA) node. Beginning at this point, the electrical signal depolarises causing the contraction of the atria and pushing blood down into the ventricle (Seeley et al., 1991).

The electrical impulse travels to the AV node. This node acts as a gate, slowing the signal 0.1 seconds down, with the view to the atria and ventricles do not contract at the same instant. It is very important because otherwise, they would be pushing against each other and blood would not be able to move through the heart in a coordinated way.

Special fibers called Purkinje fibers load the signal through the walls of the ventricles, the ventricular depolarization begins and at the same time, the atria repolarise. The ventricular repolarization occurs after the ventricular contraction.

This process is a continuous cycle. Since it is electrical, we can measure it using electrodes in

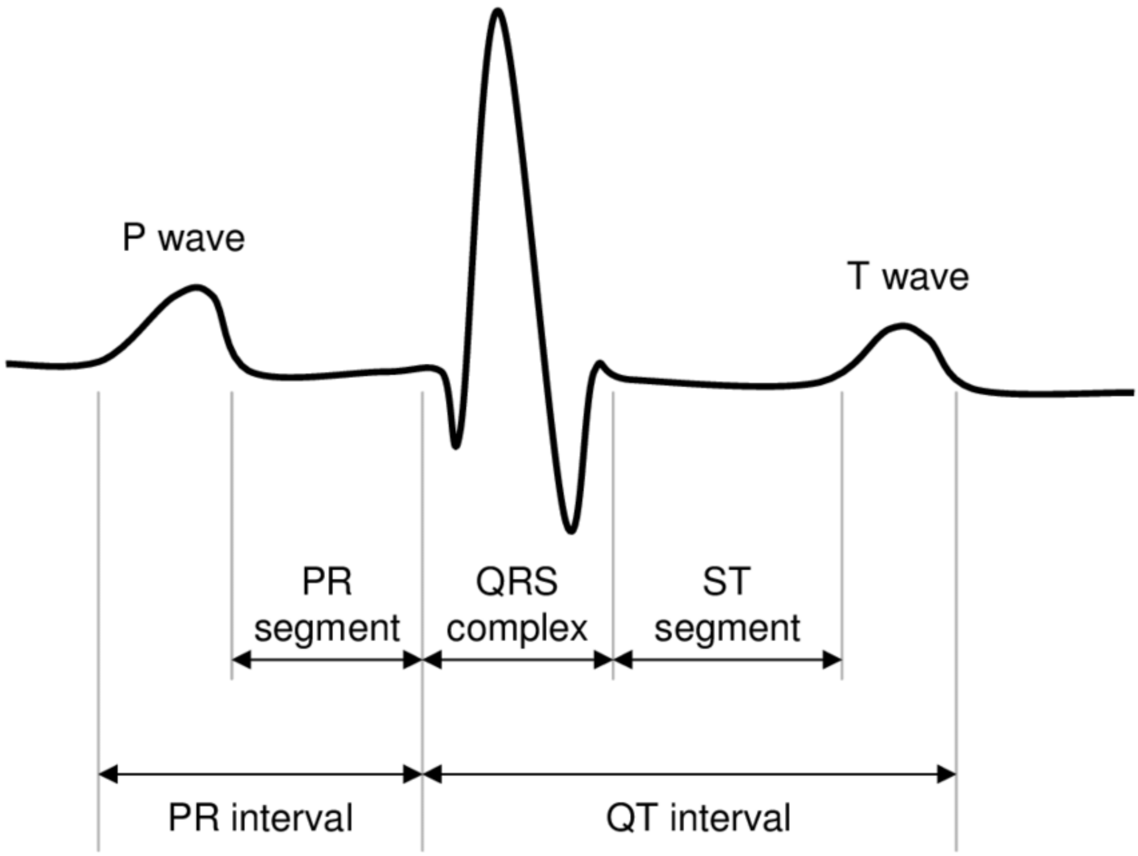


Figure 2.2: Single Heartbeat (Malasri and Wang, 2009).

the body and we obtain ECG, and in ideal acquisition conditions we can extract some important parts: the P, Q, R, S and T waves, each corresponding to a single heartbeat as can be seen in Fig. 2.2 (Seeley et al., 1991).

2.2 Acquisition

As mentioned before, to study the signal we need strategies to collect the ECG from the patient. In this section, we will cover the standard medical acquisition and the acquisition to biometric recognition.

2.2.1 Standard Medical Acquisition

2.2.1.1 12-Lead Configuration

We can consider the heart as a dipole set up in the middle of a sphere constituted by the human body, where the arms and the left leg are approximated to the vertices of an equilateral triangle. As such, the vectors of electric conduction also form an equilateral triangle (see Fig. 2.3). Thus, the tensions measured in each derivation are proportional to the projections of the electric vector

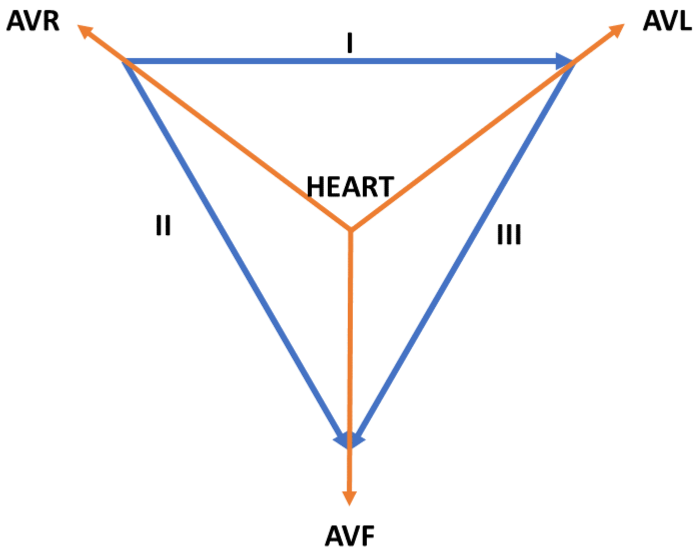


Figure 2.3: Einthoven Triangle[1].

of the heart on each side of the triangle (see Eq. 2.1) (Lin and Sriyudthsak, 2016) (Bronzino and Company, 2000).

Wilson proposed a way to make unipolar measurements of potential. Ideally, these would be measured against an infinite reference. Thus, the central terminal consists of connecting leads through the same resistor value to a common node. The voltage at this Wilson central node corresponds to the mean of the voltages in each branch. Normally, 5 MΩ resistors are used, but the higher the value of the resistors, lower the common-mode and the sizes of artifacts introduced, by the electrode-skin impedance. Thus, it is seen that the central terminal of Wilson is in the center of the triangle of Einthoven (see Fig. 2.3).

We can modify the leads for increased leads if we remove the link between the member to be measured and the center terminal, which results in a 50% signal amplitude increase. These leads are known as aVL, aVR, and aVF. In addition, 6 leads were introduced to measure potentials near the heart, making a total of 12 leads (see Fig. 2.4). They are grouped into three categories: six monopolar precordial leads, three monopolar limb leads, and three bipolar limb leads. The limb leads allow us to get signals from the frontal plane and the precordial leads on the axial plane (Bronzino and Company, 2000).

$$V1 - V2 + V3 = 0 \tag{2.1}$$

[1] http://en.my-ekg.com/basic-principles/leads-ekg.html

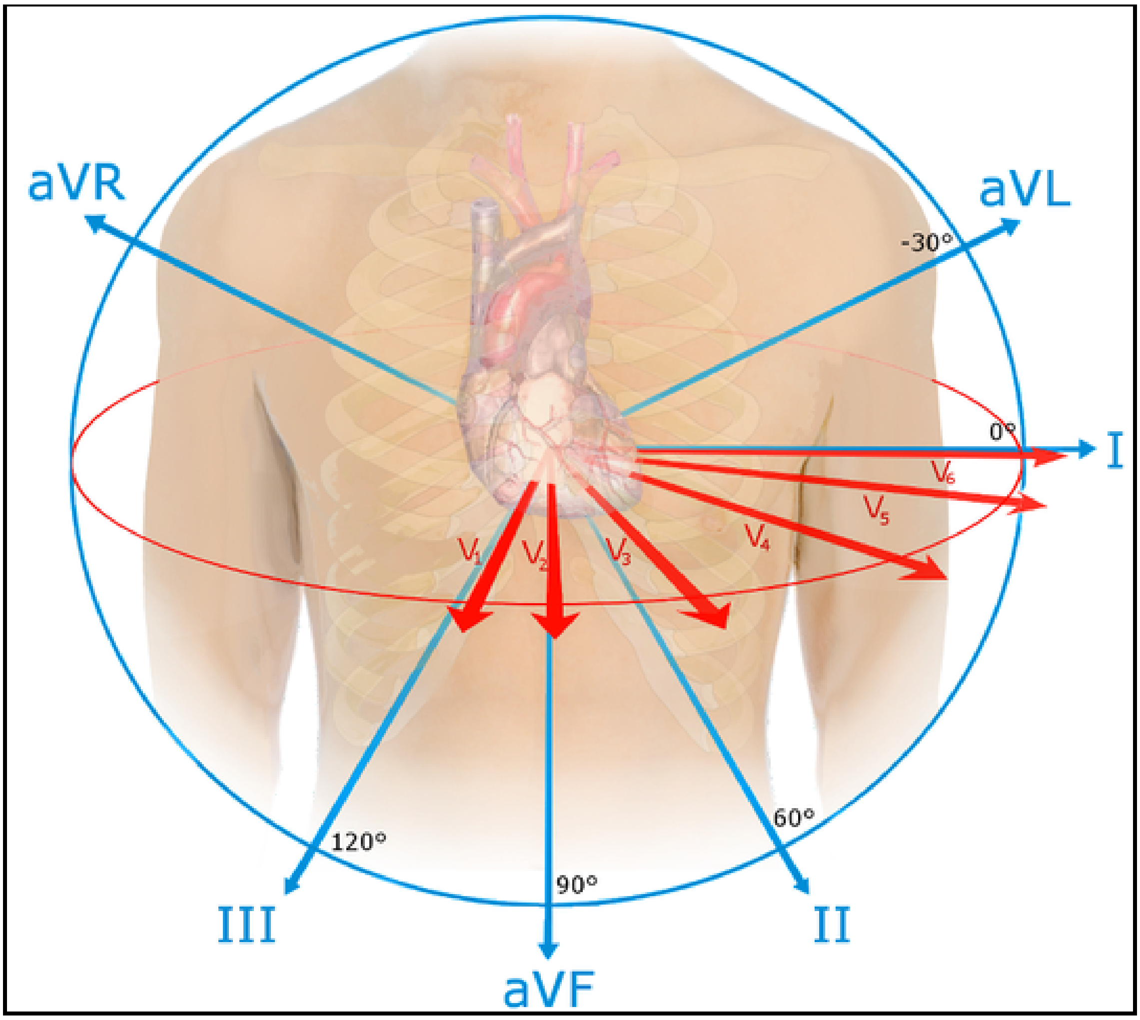


Figure 2.4: 12 Leads Configuration[2].

2.2.1.2 Corrected orthogonal configuration (Vectorcardiography Lead System or Frank VCG systems)

Frank's lead system was used, in the beginning, for patients in the supine position, with electrodes placed in the fourth intercostal space. Eventually, there was the need to decrease the error due to inter-individual variation (human anatomy), as so it was increased the number of electrodes to seven. It is used scalar leads X, Y, and Z to form the orthogonal lead system. Unlike the 12 leads configuration that uses the frontal and the axis plane (Bronzino and Company, 2000).

Normally the electrodes are placed on the chest, specifically on the left (point A), front (point E), right (point I), and back (point M) as it is possible to see in Fig. 2.5. Close to the heart (between point A and E), point C is chosen because it is easier to obtain a better signal. In addition, to have a better ground truth and noise removal it is included a point in the neck and one on the left foot (Arrobo et al., 2014).

[2] https://commons.wikimedia.org/wiki/File%3AEKG_leads.png

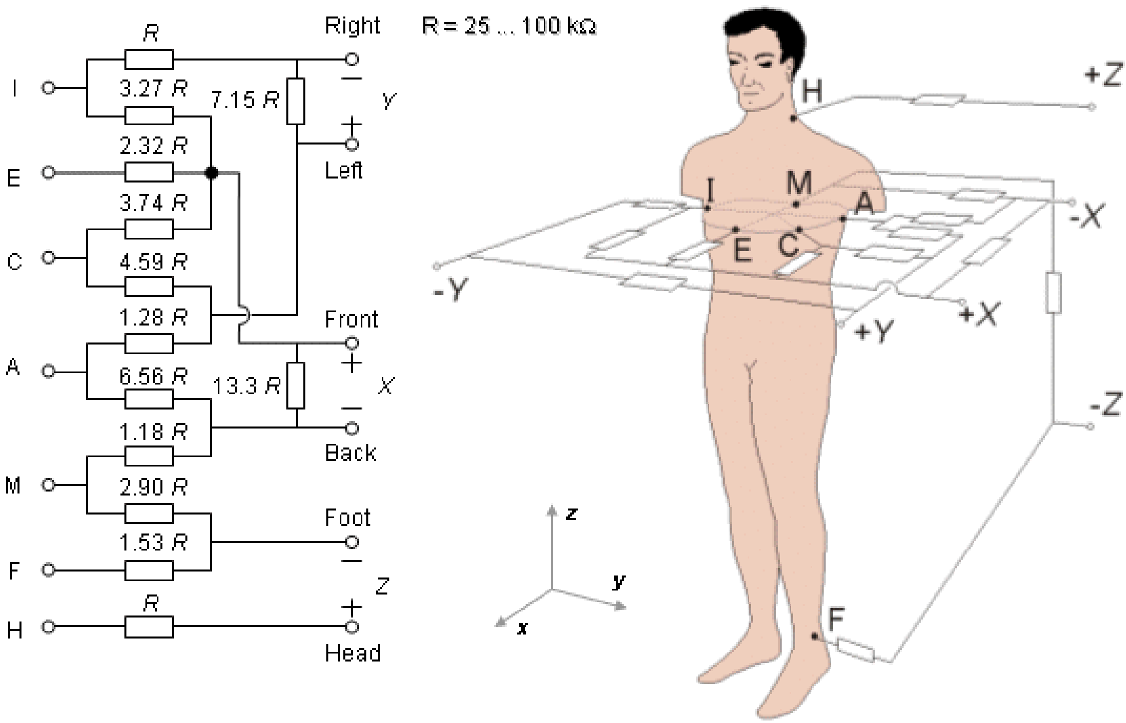


Figure 2.5: Electrode placement in the Frank VCG system[3].

2.2.2 Common ECG Acquisition Settings

2.2.2.1 In-the-Person

Devices in this category are situated inboard the human body. Normally are used only in clinical scenarios to monitor medical conditions. The most widely known example is artificial pacemakers (implantable devices), used to compensate the malfunction of the heart electrical system (da Silva et al., 2015) (Merone et al., 2017).

2.2.2.2 On-the-Person

Nowadays, on-the-person approaches to ECG measurement are the most common, where the signal is obtained through body surface electrodes with the help of a device. This group of devices includes the 12 leads measurement and Frank VCG. With the evolution of technology, it is possible to measure the ECG with wearable devices, like VitalJacket from BioDevices. These approaches have also been commonly used in ECG biometric systems (da Silva et al., 2015) (Merone et al., 2017) (Cunha et al., 2010).

2.2.2.3 Off-the-Person

In this approach, the sensors are commonly small metallic electrodes to hold in the fingers with the goal of integrating into objects of the daily routine. Here, the user must wear the sensor or

[3] http://www.bem.fi/book/16/16.htm

enter voluntarily in physical contact with the sensor. In order to perform ECG acquisition easier through the hand palms or fingers da Silva et al. (2015) proposed a bipolar sensor with virtual ground and dry electrodes. Off-the-person approach allows measuring the ECG at distances of 1 cm even with clothing between the body and the sensor (da Silva et al., 2013) (Bor-Shyh Lin et al., 2013). Another examples is the steering wheel from *CardioId*[4]. Although they are used in biometric systems, it is revealed to be a hard task to remove noise.

2.2.3 Variability

In the previous subsection, we detect several factors like an electrical current which makes hard to obtain good results. In consequence of that, the final signal is perturbed.

There may be a lot of possible contamination, in the moment of the signal acquisition and that can contribute to a higher degree of variability. Also, as Labati et al. (2014) shows even in a good acquisition and in normal conditions the subject can have a high variability known as intrasubject variability in only 24h (Pummer, 2016) (Schijvenaars, 2000).

Several interfering factors exist and can influence the variability on a short-term (ST) or long-term (LT). They can be divided into five categories: environmental factors (LT), lifestyle factors (ST and LT), physiological and pathological factors (ST and LT), non modifiable factors and effects and electrode characteristics and placement (ST) (Fatisson et al., 2016) (Schijvenaars, 2000) (Pummer, 2016).

Physiological and pathological factors also contribute to variability, such as endocrine factors, respiratory factors, neurological factors, cardiovascular diseases, stress, depression as it is discussed by Fatisson et al. (2016). Others factors like body temperature and lifestyle factors such as physical activity, alcohol, tobacco, drug dconsumption, and meditation are discussed by van Ravenswaaij-Arts et al. (1993) and Fatisson et al. (2016). Environmental factors like electromagnetic fields, vibrating tools, fatigue also affect variability (Schijvenaars, 2000) .

As discussed in the literature (van Ravenswaaij-Arts et al., 1993) (Fatisson et al., 2016) (Labati et al., 2014), in biometric systems the objective is to decrease the intra-variability (the variations between heartbeats of the same person) and increase the inter-variability (the variations between heartbeats of different subjects) of the signals.

2.2.4 Noise Contamination

Besides the variability factors discussed above, the signal can also be contaminated and destabilized by noise. The higher amplitude of the heartbeat is in the QRS complex and is between 2-3 mV, so is very common to have a lot of noise in the signal.

Signal noise can have multiple sources, like:

• Baseline wander and motion artifact: Usually result of the subject movements. The baseline wander is caused by breathing movements (involuntary). The motion artifact is commonly associated with random limb movements (Singh et al., 2015).

[4] https://www.cardio-id.com/cardiowheel

• Powerline interference: Caused by variation in the electrode impedances and to stray currents through the patient and the cables. It is necessary to remove the AC component, equivalent to the noise of high frequency (consists of 50/60 Hz AC) (Singh et al., 2015) (Levkov et al., 2005).

• Muscle Contraction: Since we are measuring electrical signals and some of the involuntary muscles like the heart contract, electric impulses are generated. When detecting the ECG, it is possible to obtain the signal from the contraction of other muscles.

• Electromagnetic noise: It is induced by the surrounding electronic equipment. The protection circuit limits the maximum voltage that can enter the amplifier to minimize saturation and protect the system. It can be reduced by using magnetic shielding (difficult) or by winding the wires to decrease the area of the loop (Bronzino and Company, 2000) (Singh et al., 2015).

• Electrode movement/contact loss: Since the electrodes are in contact with the skin can cause skin irritation. Skin-electrode impedance is very important and can deteriorate the signal (Taji et al., 2014).

• Misplaced Lead: If put the lead in a wrong position that creates electrode contact noise, the ECG will be affected by a frequency of 60Hz (Taji et al., 2014) (Singh et al., 2015).

• Pacemaker interference: Like another device, the artificial pacemakers can interfere in ECG signal captured (Singh et al., 2015).

• Ground Loops: When two machines are attached to the patient, there are two grounds. This may lead to slightly different stresses, creating a current passing through the patient producing a common mode voltage on the electrocardiograph. Aside from being a safety issue, it can raise the potential of the patient (Cutmore, 1999).

Since biometrics requires off-the-person measures and as mentioned before in subsection 2.2.2.3, the noise comes up naturally because the environment is less controlled. That can lead to a high signal to noise ratio.

2.2.5 Right Leg Drive

In most modern electrocardiographs, the patient is not connected to ground. Instead, the right leg electrode is connected to the output of an auxiliary operational amplifier. The common mode voltage is felt by two resistors and the signal is inverted, amplified and routed back to the leg. This negative feedback allows decreasing the common-mode voltage, which reduces interference on the electrocardiograph. It can also provide some electrical safety. If any unusual high voltage appears, the auxiliary operational amplifier saturates, and the patient is "disconnected" from the ground because the amplifier can no longer consume for the right leg, and there is no current passing, protecting it (Bronzino and Company, 2000) (Kesto, 2013).

Wang et al. (2011) applied the right leg drive to the signal in order to reduce common mode noise, allowing to improve the quality of the signal.

In the literature in ECG-based biometric recognition, in the off-the-person acquisition normally the right leg drive is not applied.

2.3 Biometrics Application: Challenges and Opportunity

At this point, we are aware that electrocardiogram measurement depends on many factors. During this dissertation will be used on-the-person data. Future approaches intended to use off-the-person that increase the challenges in ECG biometric systems.

A biometric system needs to be effortless for the user. That provides sustainable to the off-the-person acquisition, but the obstacles like variability and noise described before can difficult the task.

Since the electrocardiogram is a continuous and cyclic signal we will cover the continuous ECG-based biometric recognition, especially techniques of template update will be studied. In the next chapter, the fundamentals of machine learning are presented.

Chapter 3

Machine Learning: Fundamental Concepts

3.1 Introduction

Machine Learning (ML) is a field of artificial intelligence that learn from given data in order to recognize patterns. Aims in predict a variable Y from features extracted from data, X, through a function.

In this chapter, the basics of machine learning will be presented including supervised, unsupervised and reinforcement learning, feature extraction, dimensionality reduction, classification and performance evaluation.

3.2 Supervised vs Unsupervised vs Reinforcement Learning

One of the tasks in machine learning is *supervised learning*, where the training data comprises the input vector together with target vectors. A classification problem happens when the target vector is a set of a finite number of discrete categories. We call regression when the output is one or more continuous variables (Bishop, 2006).

Another task in machine learning is *unsupervised learning*, where we do not have targets. The system learns with unlabelled data, identifies likeness data and decide based on the presence or absence of the similarity in each new piece of data (*clustering*, groups with similar data). We can estimate the data distribution or design the multidimensional data in $2D$ and $3D$ for visualization (Bishop, 2006).

Finally, *reinforcement learning* correspond to an attempt by the agent of approximate to the environment's function (see Fig. 3.1), such that we can send actions into a black-box that stays in the environment that maximize the rewards it split out, the agents learns from the success/mistakes (Kaelbling et al., 1996). Some of the basic concepts are described below:

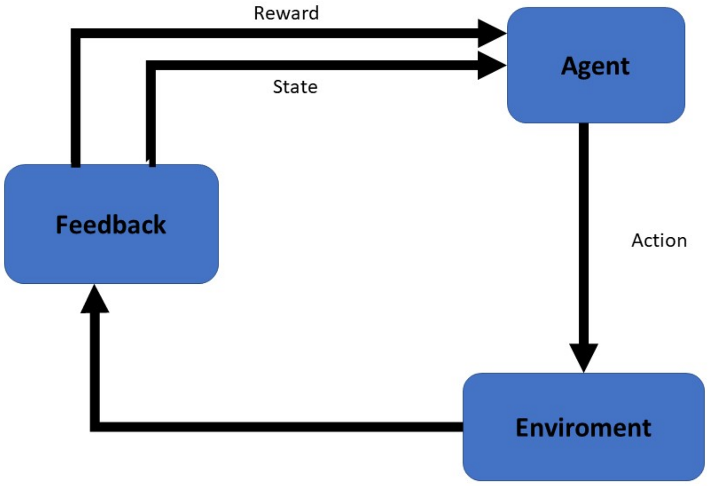


Figure 3.1: Basic modules of reinforcement learning.

• *Agent:* In this context corresponds to the algorithm and takes several *actions* (set of possibles moves that the agent can do) according to the *policy* established (defines the behaviour of the agent). Does not know the function of the environment (Kaelbling et al., 1996).

• *Environment:* Corresponds to the surrounding world where the agents take actions. The functions are what transform an action taken in the current *state* (the immediate situation that the agent finds itself) into the next state and a *reward* (feedback that allows us to measure the success or failure of an agent's action, it can be immediate or delayed) (Kaelbling et al., 1996).

• *Discount factor:* Correspond to a factor that converts future rewards worth less than immediate rewards. (Kaelbling et al., 1996).

3.3 Feature Extraction

One of the main steps in machine learning is the feature engineering (begins with a piece of informative data and will organize them) because is with these that all the system will work. We can consider feature extraction a dimensionality reduction process (will be explained forward) because we will reduce the data in a more manageable group also known as feature set (Guyon, 2006). As mentioned by Guyon (2006), the pre-processing transformations may include: normalization, signal enhancement, extraction of local features, standardization, linear and non-linear space embedding methods, non-linear expansions and feature discretization.

3.3.1 Feature Extraction in ECG Biometric

In ECG biometrics, the feature extraction approaches are generally grouped into two categories: *fiducial* and *non-fiducial*. *Fiducial approaches* are very commonly used in biometric systems that make an on-the-person signal acquisition. Fiducial points are landmarks on the ECG complex such as the baseline (PQ junction), and the onset of individual waves such as PQRST. Normally the P wave, QRS complex and T wave are find out and from them are extracted several features as can be seen in Fig. 3.2. That means that fiducial point-based methods need the precise boundaries of the waveforms (Plataniotis et al., 2006) (Israel et al., 2005) (Biel et al., 2001).

The most common features described in the literature are:

• The amount of elapsed time between events such as P duration (time of P wave), ST duration (time from S wave until T wave), QT duration (time from Q wave until T wave), QS duration (time from Q wave until S wave), QRS duration (time of the complex QRS), PQ interval (time from P wave until Q wave), RS interval (time from R wave until S wave), PT interval (time from P wave until T wave), RT interval (time from R wave until T wave);

• The maximum distance measured from a position of a vibration or oscillation, such as RS amplitude (distance from R wave until S wave), ST amplitude (distance from S wave until T wave), QR amplitude (distance from Q wave until R wave);

• The inclination of the segments such as RS slope (inclination between R wave until T wave), ST slope (inclination between S wave until T wave);

• QRS onset that correspond to the start point of the QRS wave.

Since biometric research started using off-the-person acquisitions, the signal quality is worse. That difficult the extraction of fiducial features and lead to a new methodology: *Non-fiducial approach*. Instead of extracting the individual ECG pulses, these methods consider an arbitrary window (example 5 seconds) of the Biometric trait. This kind of approach has a condition, commonly the window size are bigger than one single heart beat (Plataniotis et al., 2006) (Jung and Lee, 2017). The common methods are Autocorrelation (ACL), Discrete Cosine Transform (DCT) and Continuous Wavelet Transform (CWT).

The fiducial approach has best results in biometric systems but requires the morphological features capture of the ECG signals. The biometric system complexity increase with these features and need a medical acquisition to have a clean signal to be possible to identify all the fiducials. The non-fiducial approaches have been studied to overcome the disadvantages of fiducial methods. The biggest advantages are no need for synchronization of the heartbeat pulse, and no need of exact heart rate detection, since it differs on each record data and also varies over time (Plataniotis et al., 2006) (Jung and Lee, 2017) (Tan and Perkowski, 2017).

3.4 Dimensionality Reduction

After feature extraction, the number of features can become too high for a time-efficient recognition, because it demands a large amount of memory and computation power, or it can ease overfit.

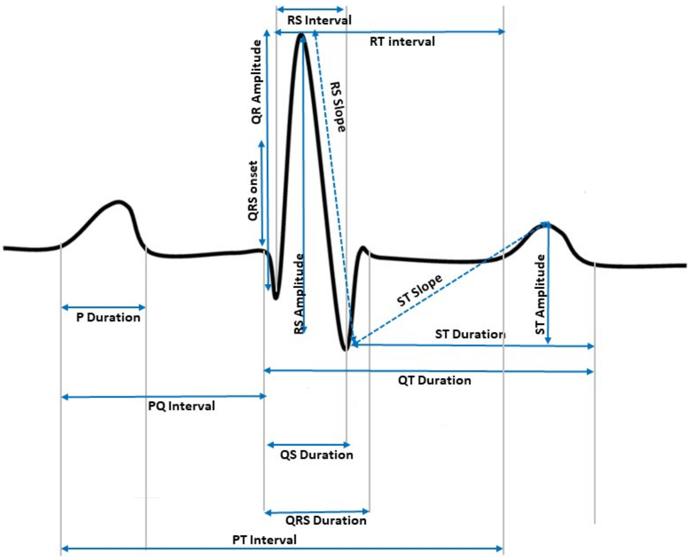


Figure 3.2: Some of the most common fiducial features extracted in ECG biometrics.

Dimensionality reduction is used to keep the features with the maximum discriminant power and reduce the computational cost for the purpose of improve the performance. Also, can be useful for outlier removal since reduce some features that can contain out range values. Every feature selection method follows these characteristics: research direction, research strategy, evaluation strategy, selection criterion and stopping criterion (el Ouardighi et al., 2007).

Several methods exist such as:

• ANOVA: The one-way analysis of variance is normally applied to determine differences between two or more independent groups, through statistics. The selection of features is performed according to the highest scores (Lee et al., 2013).

• RFECV: The recursive feature selection cross validation determine in each iteration, the worst N features, eliminating them. The remaining features that gives the maximum score on the validation data, is considered to be an optimal number of features (Yeoh et al., 2017).

• SFS: Sequential forward selection is an iterative algorithm that starts with the best individual feature, and adds, at each iteration to the selected subset, the feature that maximizes the criterion function (Zhang and Jain, 2004).

• CFS: Correlation feature selection makes the assessment of subsets of features using a criterion function that evaluates the correlation of the features with the classification, yet uncorrelated with each other (Zhang and Jain, 2004) (Hira and Gillies, 2015).

• PCA: The Principal Component Analysis is a subjective approach for dimensionality reduction because seeks a projection that best represents the data by finding a linear transformation that preserves best the data. It is used to emphasize variation and bring out strong patterns in a dataset, removing correlated features (Zhang and Jain, 2004).

• KPCA: The Kernel Principal Component Analysis is an extension of PCA but with kernel methods. Basically finds a hyperplane that divides the points into arbitrary clusters, but while PCA is confined to linear transformations, KPCA can find a non-linear manifold (the data are spread to a upper dimensional feature space) which is non-linearly related to the input space (Widjaja et al., 2012).

• DCT: The Discrete Cosine Transform is an appropriate and flexible choice for a data compression algorithm that eliminates less significant features. This method divides the signal into smaller parts and some of the coefficients are selected to construct significant feature vectors or ECG biometrics recognition (Allen and Belina, 1992).

• Wilkes lambda stepwise correlation: It is a particular case of ANOVA and measures the individual discriminative power of the variable and as smaller Wilkes Lambda, is more discriminative. (el Ouardighi et al., 2007).

• LDA: The Linear Discriminant Analysis is used when the measurements made on independent variables for each observation are continuous values. Similar to PCA, but LDA endeavours to model the variation between classes of data, and so it requires prior labelling (which can be a disadvantage) (Song et al., 2010) (Karafiat and Burget, 2005).

• HLDA: The Heteroscedastic Linear Discriminant Analysis allows to preserve useful dimensions and separate the features vectors that represent individual classes. Differs of LDA because requires the estimation of the covariance matrix for each class (Karafiat and Burget, 2005).

• SHLDA: Smoothed HLDA similar to LDA and HLDA. Differs from HLDA only in the estimation of the covariance matrices class. The estimation depends on the smoothing factor that goes in the interval of [0,1] (when it is 0 SHLDA behaves like LDA and when is 1 becomes HLDA) (Karafiat and Burget, 2005).

• FLDA: The Fisher Linear Discriminant Analysis evaluates locally the levels on the between-class scatter (interclass) and the within-class scatter (intraclass) (Sugiyama, 2007).

• GBFS: Greedy Best First Search always selects a state with minimum which heuristic value among all candidates (Heusner et al., 2018).

3.5 Classification Methods

To have a score to classify our system we need a classifier to train our model. Normally we have a train set with several samples, and we divide this train group into two parts: one for training and other for testing. Inside of the learning step, we can have a training set (learn the model) and a validation set (used to tune the model's hyper-parameters). We also can have cross-validation, that is used when we have a few numbers of samples and a proper validation set does not exist. Here, we take the data and split in K number of folds. In 5-fold cross-validation we train the model in 4 dataset partitions and test in the other. In leave-one-out, we train in N-1 samples and test in the other and repeat the process to each sample (only used if the number of samples is too low) (Costaridou, 2005) (Beutel, 2000). This technique allows us to have a better understanding of the real performance of the system.

There are several different classification methods, including:

• Naive Bayes: Having a training set, we model each class distribution through normal distribution. Given a new point, we compared the distance to its distribution with the Mahalanobis distance (Euclidean distance if the matrix was unitary). The nearest distribution is the one that most likely generated the new point, and we assign this class to it (Hastie et al., 2009) (Raschka and Mirjalili, 2017).

• Logistic Regression: Is a classifier that generates a coefficient that maximizes the similarity of observing the samples values. Tries to look for the best model that describes the relationship between variables, the outcome (dependent) and the predictor/explanatory (independent). The typically variable hyper-parameter is the cost (James et al., 2013).

• Decision Trees: Is an algorithm created to solve a classification problem, by doing several "questions" regarding the attributes of the test record (Hastie et al., 2009).

• Random Forest: Are an extension of Decision Trees. Creates a forest, a group of several decision trees. Normally, how much more trees in the forest, more robust the classifier looks like. Parameters as depth, maximum number of leaves, maximum leaf nodes can be found in these two algorithms (Hastie et al., 2009).

• Multilayer Perceptron (MLP): Is a class of a feed-forward artificial neural networks, with nodes that use non-linear activation functions. The tunable hyper-parameters chosen are the activation function, alpha (which is a regularization parameter that penalizes weights with large magnitudes), the hidden layer size and the learning rate (Raschka and Mirjalili, 2017).

• Nearest Neighbour: Is a non-parametric method that aims in search similarity on vicinity instances. Does not require a model and uses the data directly for classification. The number of neighbours is the hyper-parameter that need to be choose. Normally is odd to avoid ties (James et al., 2013). The kNN can have several different metrics for classification such as Euclidean, Cosine, and Mahalanobis distance. For example Guennoun et al. (2009) used the Mahalanobis distance for matching each heartbeat with the template stored in the system database. According to a threshold, decisions are made.

It's a method that is simple to implement, "training" is very efficient, adapts well to online learning, robust to noisy data but is sensitive to feature value ranges, the classification can be time-consuming and the memory requirements are high.

• Support Vector Machines (SVM): Performs classification in order to find a linear solution (hyperplane) to separate two classes and maximizes the margin between them. The vectors between them that define the hyperplane are the support vectors. In the case of the problem is not linearly separable we can resort to the *kernel trick* that consist of a transformation of the problem in a bigger dimension (Raschka and Mirjalili, 2017).

• Bootstrap Aggregating: Generates weak classifiers/predictors. To make a decision it is used the aggregated average of weak classifiers. Good to use when we have an unstable classifier prediction like in ECG heartbeats data (a little shift in the training data can lead to a big exchange in the construction of the classifier, which will lead to a change in accuracy) (Louis et al., 2016).

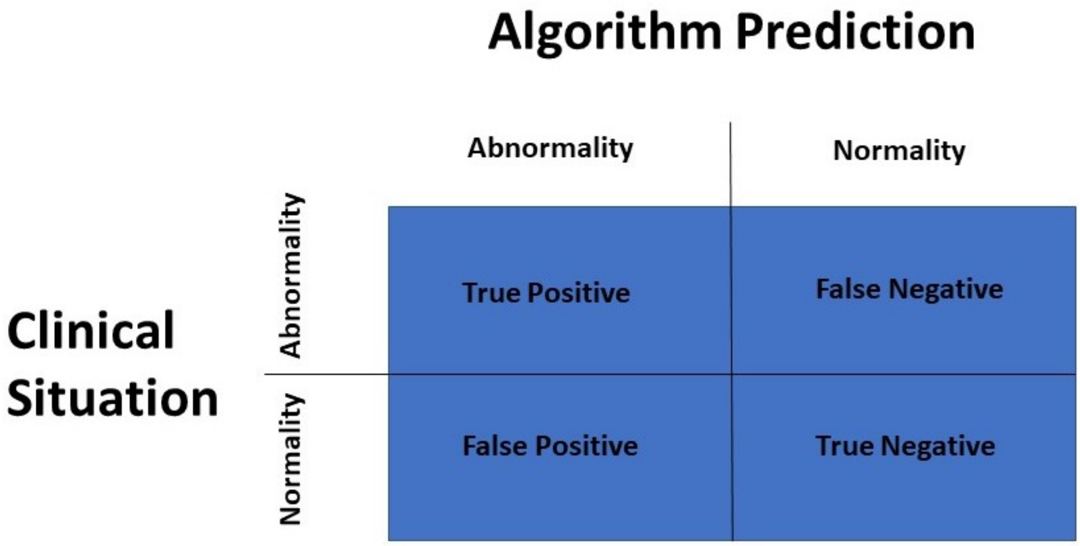


Figure 3.3: Example of a confusion matrix for a binary medical classifier

3.6 Performance Evaluation: Basic Metrics

Regarding the performance of the system, we need to use some metrics. To better understand the most relevant error metrics, it is needed to understand the basic ones such as TP (true positive), TN (true negative), FP (false positive), FN (false negative) as can be seen in Fig. 3.3 (Sim et al., 2007) (Costaridou, 2005) (Beutel, 2000)

• Mean error: When the goal is to estimate the value of some scalar or vector quantity. We can focus on the absolute or relative difference between the calculated result and some independent measure of the true value (see Eq. 3.1).

$$MRE = \frac{|\text{ measure value} - \text{true value }|}{\text{true value}} \tag{3.1}$$

• Accuracy (ACC): Measure the correct decisions made by algorithm (see Eq. 3.2). Accuracy may not be a useful measure in the case where there is a large class skew or in case of detection because in this case did not interest to hit in the negative (eg 99% of accuracy with 98% of negative instances).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.2}$$

• Sensitivity: Concerns how frequently the algorithm classify positive being the value real and positive (see Eq. 3.3). Is also called true positive fraction (TPF) or Recall.

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.3}$$

• Specificity: Contrasting, concerns to how often the algorithm classify as negative when the values are real negative (see Eq. 3.4). False positive fraction (FPF) is the same as $(1-specificity)$. All these measures can be stated as a fraction between 0 and 1, or as a percentage between 0 and 100%.

$$Specificity = \frac{TN}{TN+FP} \tag{3.4}$$

• Precision: The fraction of detections that are relevant (see Eq. 3.5).

$$P = \frac{TP}{TP+FP} \tag{3.5}$$

• F-measure: Match precision and recall. The traditional measure is the F1 measure (see Eq. 3.6), where P and R are weighted equally. F2 is also commonly used as a measure (see Eq. 3.7), that consider recall weights double comparatively to precision, and F0.5, which precision weights twice as much as the recall.

$$F1 = \frac{2 \times P \times R}{P+R} \tag{3.6}$$

$$F\flat = \frac{(1+\flat^2) \times P \times R}{R+P \times \flat^2} \tag{3.7}$$

• ROC-Curve: Can be described as the space of possible tradeoffs between sensitivity and specificity. The ROC curve is commonly plotted (see Fig. 3.4) with TPF (Sensitivity) on Y-axis and the FPF (1-specificity) on the X-axis. Afterwards, the algorithm runs with different values for this parameter to obtain several TPF/FPF pairs. The ideal performing point is the upper left corner, with TPF = 1 and FPF = 0.

• AUC: Generally it is used the area under the ROC curve (see Fig 3.4) as a efficiency measure for evaluating the quality of the curve (consequently the algorithm that produced the ROC curve). The higher the AUC, the greater the probability of making a appropriate decision. One algorithm is globally better than another algorithm if its ROC curve has a greater AUC. However, since the AUC is an overall measure, it may hide important details.

• Free Response ROC-Curve: It is appropriated for when the target is to identify all positive instances in an image. Correspond to have TPF in one axis and FP/image in another axis (for a given value of TP how much FP exist in mean per image). Increasing the sensitivity it increases the number of FP. It is also used the AUC as metric of evaluation.

• Hausdorff distance: Defined as the maximum of the minimum of two distances between two contours of the same object. It is not a true distance metric because it does not verify the commutative property.

• Bland Altman plot: Also called difference plot, it is used to compare two quantitative measurements (two different parameters that measure the same property, should be correlated). It is a method to quantify the similarity between two quantitative measurements by constructing limits of agreement.

• Koppa statistic: Also called Cohen's Kappa, it evaluates the agreement between observers (the inter-observer agreement). It applies in medicine to give a quantitative measure of agreement

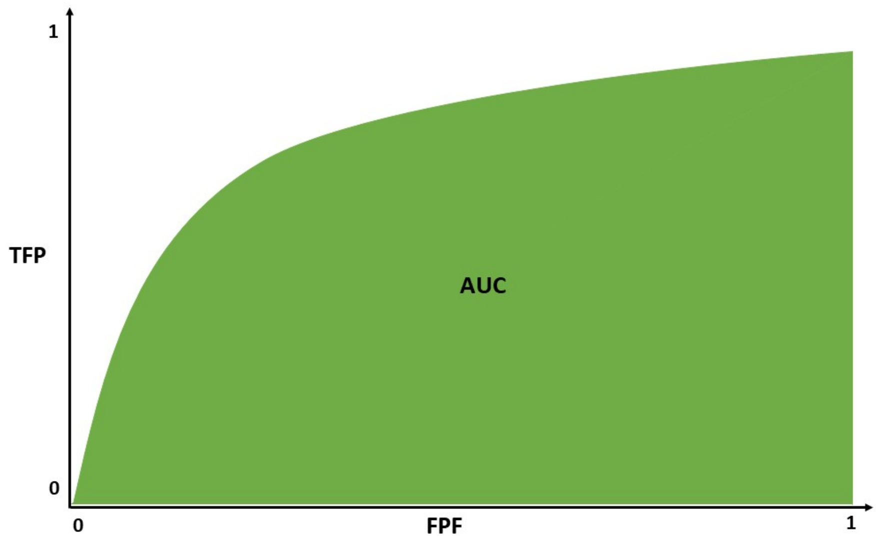


Figure 3.4: ROC curve with AUC in authentication problems, FPF (false positive fraction) and TPF (true positive fraction)

between ratters, such as in physical exam findings, segmentation, pathology grading, among others. While the Bland-Altman plot compares the agreement on real value measurement, Kappa compares categorical labels (see Eq. 3.8). Kappa varies in the interval [-1, +1], where 1 is perfect agreement, 0 corresponds to a selection by chance, and the correspondent negative values reveal less agreement than randomly, i.e, high disagreement between the observers:

$$K = \frac{po - pe}{1 - pe}, \tag{3.8}$$

where p0 is the probability of full agreement and pe is the expected agreement.

3.7 Conclusion

Every aspect discussed in this chapter is very important in biometrics systems since it requires the extraction of the most important features, selection of the best of them and the choice of the model that adapt better to our data in order to have a higher score.

Machine learning techniques are diverse and each one has specific advantages and can be a powerful tool to biometric systems. In this dissertation, we will often resort to machine learning functions for the classification process.

Chapter 4

Biometric Systems: Basics

4.1 Biometric Modalities

Since the first proposal of a biometric system, some different modalities were proposed over time, with the objective of proving the best methodology to guarantee a unique human identification and authentication. Prabhakar et al. (2003), Jain and Kumar (2012) and Delac and Grgic (2004) have listed the most common biometric modalities, like fingerprint, face recognition, hand geometry, iris, voice, palmprint, signature, hand veins, gait, keystroke, and others less common such as ear shape, odour, DNA. Some of these, are ready to commercialize, like a fingerprint which is a very common method in the last generation of smartphones. However, with the evolution of technology, new techniques of counterfeit are used, and already have some successfully spoofing attacks in this area (Hadid et al., 2015), (Akhtar et al., 2015).

Some fingerprint replica can be made using a mould of silicon, iris can be spoofed using contact lenses, voice can be recorded from a phone, faces can be replicated from a 2D photo, DNA can be stooled (Akhtar et al., 2015).

To avoid this, some new approaches were proposed. These kinds of approaches normally guarantee liveness detection, in order to confer resistance to spoofing methods (assure that is a person who is there and not a mold).

Agrafioti et al. (2012) suggest that some of these vital signals carry information that is unique for each human like the electrocardiogram (ECG), phonocardiogram (PCG), photoplethysmogram (PPG), blood volume pressure (BVP), and heart rate variability (HRV). The ECG biometric furnish inherent liveness detection which has computational benefits. Like other vital signals, the ECG signal is very difficult to collect for spoofing purposes (Eberz et al., 2017a).

Nevertheless, vital signals present some disadvantages, and the most significant is their variability, due to several conditions. One good example of this kind of problems is found in the ECG: it can change in a short period of time as shown by Labati et al. (2014).

Nowadays, another approach that is beginning to be implemented is the hybrid system, which combines biometrics traits with traditional credentials like Multi-biometric Authentication System (Fingerprint + Face + iris + Password) from Raviraj Technologies.

The next step consists in combining several biometric systems (multimodal biometric systems) like Be et al. (2015) who used ECG and fingerprint.

4.2 Qualities of Biometric Modality

In the previous section, several biometric traits were presented. According to Prabhakar et al. (2003), Jain et al. (2004), Abo-Zahhad et al. (2014) and Delac and Grgic (2004) they need to fulfill these characteristics:

• Universality: Every subject must have the characteristic.

• Distinctiveness: The characteristic of two-subject must be distinct.

• Permanence: The characteristic should be sufficiently invariant over time.

• Collectability: The characteristic should be quantitatively measurable.

• Circumvention: The characteristic should be hard to mimic or counterfeit.

• Measurability: The person needs to feel comfortable when the signal is acquired, that means that needs to be an easy and quick process.

• Performance: Practicable recognition accuracy and speed.

• Acceptability: Indicates how much people are willing to accept the use of a biometric identifier.

These qualities are ideal but not mandatory. In fact, there is no biometric feature that fulfill completely these requirements.

In this work. we will use ECG as a trait, and as we can see through Tab. 4.1, it is a trait very balanced.

4.3 Common Structure of a Biometric System:

A biometric system recognizes the individual based on a feature vector derived from their specific characteristic (physiological or behavioural). Basically consist in a pattern recognition system. That feature vector is commonly designated as templates and is usually stored in a database after being extracted as can be seen in Fig. 4.1.

4.3.1 System Modules

According to Bolle (2011), a simple biometric system hold five basic components (represented in Fig. 4.2):

• Sensor Module: Captures the biometric trait data, usually resort to sensors;

• Feature extraction module: Verifies the acquired data and process them to extract feature vectors;

• Matching module: Compares feature vectors against those of the stored templates (from the database);

• Decision-making module: The user's identity is determined, or verify if a claimed identity is genuine (accepted) or impostor (rejected).

Table 4.1: Comparison between biometric traits regarding the seven characteristics (H - high; M - medium; L - low) (based on Prabhakar et al. (2003), Jain et al. (2004), Abo-Zahhad et al. (2014) and Delac and Grgic (2004)).

Biometric trait	**Universality**	**Distinctiveness**	**Permanence**	**Collectability**	**Performance**	**Acceptability**	**Circumvention**	**Measurability**
DNA	H	H	H	L	H	L	L	M
Ear shape	M	M	H	M	M	H	M	L
Face	H	L	M	H	L	H	H	M
Facial thermogram	H	H	L	H	M	H	L	M
Fingerprint	M	H	H	M	H	M	L	H
Hand geometry	M	M	M	H	M	M	M	H
Hand Vein	M	M	M	M	M	M	L	M
Iris	H	H	H	M	H	L	L	L
Palm print	M	H	H	M	H	M	M	H
Retina	H	H	M	L	H	L	L	L
Heart sounds	H	M	L	L	L	M	L	M
ECG	H	H	L	M	H	M	L	H
Gait	M	L	L	H	L	H	M	M
Keystroke	L	L	L	M	L	M	M	H
Signature	L	L	L	H	L	H	H	M
Voice	M	L	M	L	L	M	H	H

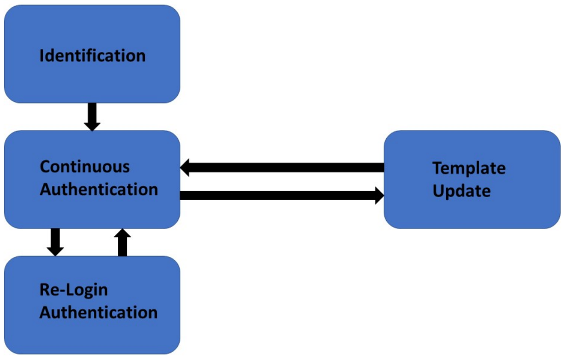


Figure 4.1: Simplest scheme of biometric for a continuous system that use template update.

• Database module: Stores the biometric data. Also, it can be updated after verification in a continuous biometric system as described by Labati et al. (2013) and Sufi and Khalil (2011).

4.3.2 Operation Modes

Two modes of operation exist in Biometric systems: verification (authentication) or identification. Concerning identification, the acquired biometric information is compared against templates corresponding to the totality of users in the database (one-vs-all problem), verification involves verify the similarity between templates that correspond to the claimed identity (two class problem, one-vs-one comparison). Hereupon, it is implies that we should deal with this two problems (identification and verification) separately.

4.3.3 Conventional vs Continuous Biometrics

Regarding the moments of identification or verification, there are two types of biometric systems. Conventional Biometrics are those that control access to a safe room/system, usually not requiring the user to re-authenticate himself for continued access (at least not for a considerable long period of time) to the protected resource, (i.e one-time verification system). In high-security environments, this may not be sufficient because the protected resource needs to be continuously monitored for unauthorized use (Sim et al., 2007) (Niinuma and K. Jain, 2010).

To solve this problem, continuous biometrics systems can be used but they are still in development phase. According to Sim et al. (2007) the continuous verification needs to follow six criteria:

1) Reliability: Any continuous system must take into account the trusty of each modality;

2) Template update: Older observations increase uncertainty, so it needs to have low weight when used to verify the legitimate user;

3) Liveness: whether the legitimate user is present or not;

4) Usability: The system must not demand re-authentication when for example the user takes a break to get some air, it needs to be convenient;

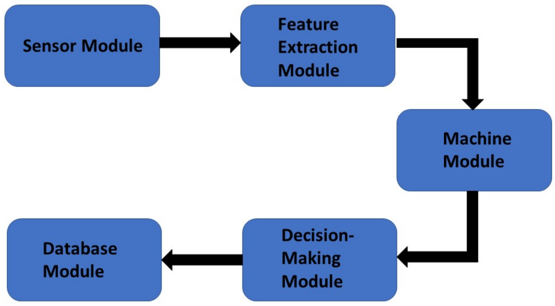


Figure 4.2: Module representation of the five basic components.

5) Security: If the user get away from the device the system must require active re-authentication;

6) Cost: The system should be design/implemented with commercial purpose (off-the-shelf devices only). Should avoid the use of special/expensive devices.

Since new data is constantly acquired we need to have in mind that not all biometric traits are viable for this process. With this in mind and with the new technologies of ECG signal acquisition (discussed in section 2.2.2), we consider that ECG signal is one of the most convincing traits because we can acquire the signal easily and be comfortable for the user.

4.3.3.1 Conventional Biometrics

To apply and predict the potential of a biometric system in a real-life situation, the evaluation of its performance is needed.

Since we are focusing on problems of identification or verification, different metrics must be used (as in Ribeiro Pinto et al. (2018), Pinto et al. (2017)). In a biometric system, we can have different problems respectively to the subject, the system or database. We can have several different situations regarding a biometric system. The situation where the subject is already registered in the system and the system identifies him correctly (allowing the access). The situation where the subject is register and the system fails in identify him (rejecting the access). Another scenario is when the subject is registered, but the system allows access under the identity of another subject. Other development that can surge in the system is when the subject is not register but can have access because the system misses the identification (identifies the subject as one of the enrolled subjects). At last, we can have a situation where the subject is not enrolled and the access is denied.

According to Grother et al. (2011) exist two types of categories errors:

• Type I: When the system fails to identify the subject (genuine user) with the respectively stored template;

• Type II: When the system matches wrongly the subject with a different person's stored template (impostors).

As mentioned by Grother et al. (2011) our goal is to maximize the situations where the subject is registered and the system identify or not the subject in order to achieve the best performance possible. The solution to guarantee the best performance can be a solid template update method.

4.3.3.2 Identification Error Metrics

Costaridou (2005), Beutel (2000), Grother et al. (2011) and Gorodnichy (2009) divide the most relevant metrics in two groups. The first group focuses on situations where the subject being identified is enrolled in the system:

• True Positive Identification Rate (TPIR or Hit Rate or rank-k identification rate): It is the desired proportion of identification transactions by users registered in the system, where the subject correct identity is among the R strongest identities returned by the system (see Eq. 4.1). The value of TPIR is called rank-one accuracy;

$$TPIR(R,T,L) = \frac{\text{No.of trials where, among L, one of the strongest R predictions above T is correct}}{\text{Total number of trials}} \tag{4.1}$$

• Accuracy: It is the correct identification of subject on total samples. When we consider the strongest prediction (R = 1) of TPIR above the threshold T (see Eq. 4.2), it is reached the identification rate. Can also be refereed as IDR;

$$IDR(T,L) = TPIR(1,T,L) \tag{4.2}$$

• Reliability: When we choose R = N (number of enrolled subjects)(see Eq. 4.3);

$$Reliability = TPIR(N,T,L) \tag{4.3}$$

• False-Negative Identification Rate (FNIR): It is the opposite of TPIR, where the test examples are classified as false although they are true (Type 2 error) (see Eq. 4.4);

$$FNIR(R,T,L) = 1 - TPIR(R,T,L) \tag{4.4}$$

• Misidentification Rate: Correspond to the opposite of accuracy, when we measure the fraction of total trials where with the subject enrolled in the system, the true identity is not the system's top ranking prediction above T (see Eq. 4.5). Equivalent to FNIR with R = 1;

$$MIDR(T,L) = 1 - IDR(T,L) \tag{4.5}$$

The other metrics are focuses on situations where the subject is not enrolled in the system:

• False Positive Identification Rate (FPIR): Is the ratio where the test examples that are classified as true, although they are false (type 1 error) (see Eq. 4.6);

$$FPIR(T,L) = \frac{\text{No. of trials with unenrolled subjects where there us ine or more predictions above T}}{\text{Total number of trials}} \tag{4.6}$$

• Selectivity: Provinient of FPIR but here we consider the average number of predictions above the threshold T (see Eq. 4.7);

$$Selectivity(T,L) = \frac{\text{No. predictions above T across all trials}}{\text{Total number of trials}} \tag{4.7}$$

• Cumulative Match Characteristic (CMC) curve, plots the rank-k identification rate against k. These rates are not enough for a full picture of the system performance.

4.3.3.3 Verification Error Metrics

As presented in section 3.6 a system can have two binary errors: FP and FN (Gorodnichy, 2009). Using a biometric system on a data set, the total number of FP (also known as False Accept) and FN (also known as False Reject) are find out to compute the cumulative measurement:

• False Accept Rate (FAR): correspond to accept incorrectly an access attempt by an unauthorized user, as can be seen in Eq. 4.8;

• False Reject Rate (FRR): correspond to reject incorrectly an access attempt by an authorized user, as can be seen in Eq. 4.9;

With both, we can define a threshold, and we can measure two different metrics:

• Decision Error Tradeoff (DET) plot, corresponds to the chart of FAR vs FRR, obtained by changing the system match threshold, and allows us to extract one important measure known as Equal Error Rate (EER) that represents the operation point where FAR = FRR as can be seen in Fig. 4.3;

• Receiver Operating Characteristic (ROC) curve that plots TAR against FAR.

$$FAR(T) = \frac{\text{Number of impostor trials where the prediction score is above T}}{\text{Total number of impostor trials}} \tag{4.8}$$

$$FRR(T) = \frac{\text{Number of legitimate trials where the prediction score is below T}}{\text{Total number of legitimate trials}} \tag{4.9}$$

4.3.4 Continuous Biometrics

In the situation of conventional biometrics (i.e. not continuous, like previously described), metrics like FAR and FRR are sensible and widely accepted metrics. Nevertheless, continuous authentication provides a unique challenge because the system accumulate errors over runtime. Here, we can have two scenarios, one corresponding to the meaning of FAR and the other corresponding to the systematic FN. In the first scenario, a FAR of $x\%$ mean that this percentage of invaders never

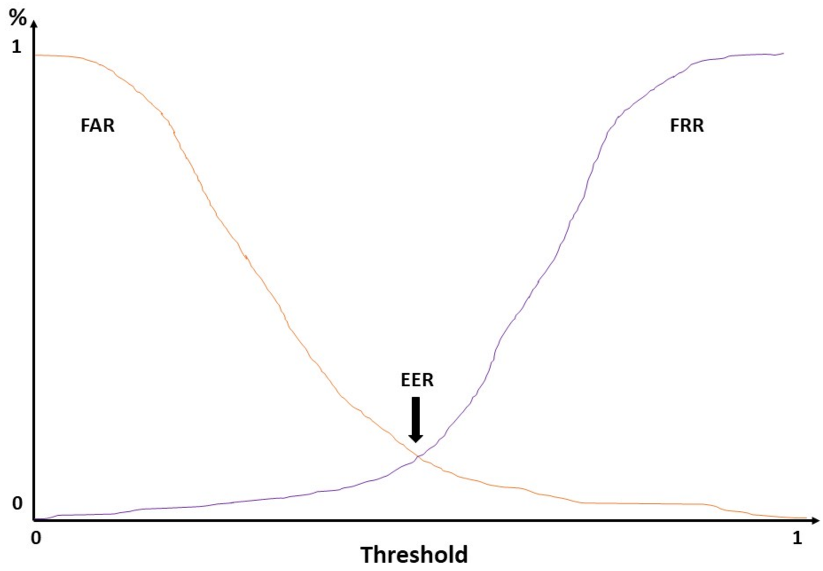


Figure 4.3: DET plot, represented by FAR (curve on green) and FRR (curve on red). EER is the intersection between the curves.

have been detected while all others are exposed immediately. In the other scenario, the systematic FN can be permanent or difficult to remove in the situation where the behaviour of the invader and the user is very similar. From a security point of view, these errors can be problematic, because an unauthorized subject can access the system and compromised it for unlimited time. (Eberz et al., 2017b).

Here we must consider the difference in the reliability of several modalities. Since the biometric trait can change over time, older biometric observations should have less weight to reflect the raised uncertainty of the continued presence of the rightful user with time. Also, the system should determine the authentication certainty in any period, even when no biometrics observations are available for one or more modalities (Niinuma and K. Jain, 2010).

Since time is not taken into account in FAR and FRR we can consider obsolete for this kind of system.

In Sim et al. (2007) they developed metrics that succinctly capture the global system performance, allowing to compare different systems.

1) Time to Correct Reject (TCR) $\rightarrow$ The authors define as the period between the beginning of the first action taken by the imposter, to the instant that the system decides to (correctly) reject him (unit is seconds). TCR should be zero (impossible computing) but if TCR fails to correctly reject the imposter, TCR will be infinite;

2) Probability of Time to Correct Reject (PTCR) $\rightarrow$ The authors define as the probability that TCR is less than the window of vulnerability (W). Ideally should be 1. The system sometimes can tolerate PTCR less than one to correctly reject an imposter;

3) Usability $\rightarrow$ The authors define as the percentage of time that the user has access to the

protected resource. Ideally, Usability should be 1, that means that the legitimate user is granted access all the time;

4) Inconvenience $\rightarrow$ The authors define as 1 – Usability. When it denied access to the legitimate user, it represents an nuisance to him because he must re-authenticate himself or take other action to restore access;

5) Average Usability $\rightarrow$ The authors define as the sum of activity-specific usabilities weighed but with the percentage of time spent on each activity;

6) Usability-Security Characteristic Curve (USC) $\rightarrow$ Corresponds to the plot Usability vs PTCR. This is analogue to the ROC-curve in one-time measure verification system. As a performance measure it is used the area under USC curve.

4.3.5 Biometric Menagerie

Each person behaves differently regarding biometric authentication systems and each of them is responsible for complicating the task of biometric authentication. In the literature it was formalized as the concept of Biometric Menagerie (R. Doddington et al., 1998) (Yager and Dunstone, 2010), (Houmani and Garcia-Salicetti, 2016) and (Teli et al., 2011), defining and labelling user groups with animal names to show their characteristics comparatively to biometric systems.

The first time R. Doddington et al. (1998) created the concept, the authors defined four types of animals that will characterize the system errors. They define the zoo as:

Sheep $\rightarrow$ The subject produces a biometric template that matches well to other biometrics templates of themselves and poorly to other people. Exhibit low FRR.

Goats $\rightarrow$ The subject produces a biometric template that matches poorly to other biometrics templates of themselves. Exhibit high FRR.

Lambs $\rightarrow$ The biometric template of a subject can easily imitate from a biometric template of a different person. Exhibit the increasing of FAR.

Wolves $\rightarrow$ The biometric template of a subject is an amazing impersonation, can imitate the other person. Exhibit a significant increase of the FAR.

Doddington's definitions did not capture the relationships between match and non-match scores. Teli et al. (2011) define a new zoo based on match scores, according to a user's relationship between the genuine and impostor match scores. They define:

Doves (sub-group of *Sheep*) $\rightarrow$ Match very well against themselves and poorly against others.

Chameleons $\rightarrow$ Generally match well with every sample.

Phantoms $\rightarrow$ Generally match poorly with every sample.

Worms (sub-group of *Goats*) $\rightarrow$ Match poorly with themselves and well with the rest of the persons.

In Yager and Dunstone (2010), the authors have made a study for comparing these two zoos' and apply this concept in an online signature. In Paone and Flynn (2011), the authors apply these ideas in iris matches for testing the consistency. As Poh et al. (2006) have done, this kind of approach is good for testing to assess user-dependent variabilities.

The zoos have in account the difference in the population and allow us to evaluate the distribution of people in different categories. That allows a better characterization of the results.

4.3.6 Ideal Conditions for a thorough Performance Assessment

After this review of the most common methodology, we realize that the most important metrics are time-dependency. Also, it is possible to conclude that we need to separate the problem in two big parts and deal with them individually. For identification, the most common metrics are accuracy or Identification Rate. For authentication, they use more FAR, FRR and EER.

Biometric menagerie appears to be influential on the system's performance, and from the way that we evaluate the system, as previously described.

4.4 System Design Considerations and Concerns

Despite all the concerns referred, when we are designing a new biometric system we need to have in account several aspects, even when we are planning a business case as stated Yanushkevich and Shmerko (2009), Wayman (2005).

• One important aspect is the business plan, where it is taken into account, with a high level of detail, the fundamental pieces for the set-up. Since is a biometric system is namely for security purpose and requires time, money and energy for the set-up;

• In real life situation the systems are not successful every time. As such, not all people will be able to use them. This implies that backup systems for exception handling will always be required;

• Despite the high level of acceptance of biometric technology will always exist some people that will reject and argue against the technology. Whereby we should study very well the needs of the population;

• Since biometric systems are commonly used for safety/security, the integration of data will be a hard task because all requirements of a biometric system must be attended by hardware/software. In order to accomplish that it must be required the authorization by the users to store the data;

• Since we are in an era of quick technology development, the product and the competitors can be in a continuous flow. So the technology that we invest today may not make sense in the next year;

• The addition/substitution of a component will inevitably lead to a change in the business process. This kind of situation must be avoided;

• Intellectual property and approval by the competent authorities are of extreme importance in order to be able to commercialize the product and protect them (e.g FDA, INFRAMED, CE) (Rossi and Lueje, 2002), (ISO, 2008).

4.5 Summary and Conclusions

The potential of Biometric systems is huge, and clearly brings a lot of advantages comparing to the traditional techniques like ID-cards as we mentioned before. To use biometric identity we require the intrinsic characteristic of the individual and guarantee the relation between the identity and person requesting access.

In this work we will focus on long term Electrocardiogram, to improve these measures and allow future comparison between the several methods (Bolle, 2011).

Nowadays, some advanced systems already use biometric modalities like face and fingerprints but as we said in this chapter, this can be manipulated and can be copied. Electrocardiogram appears as one of the best biometrics modalities because of his advantages.

However, in the recent studies, the ECG advantages did not translate in accuracy when we talk about continuous measures and that give us a space to improve this system.

Chapter 5

Adaptive ECG Biometrics Systems: Prior Art

5.1 Introduction

This chapter addresses the state-of-the-art adaptive ECG biometric systems. As so, for testing the ideas a previous knowledge regarding the significance of a well-structured signal database and the existence of long-term ECG databases is needed.

5.2 The significance of a well-structured Signal Database

The existence of a well-structured signal collection can have an impact on the machine learning problems, basically in all situations that need an organized data like our case, to guarantee reliable ECG-biometric system performance.

The significance goes further, throughout all phases of its development. In order to have the best development, prediction and performance for a real-life system, it is required having a well-structured signal collection. To accomplish such a complete collection, some considerations must be in account because have an impact on the performance of an ECG-based biometric, like the number and placement of electrodes, the sampling frequency of the acquisition, the number, posture, health condition and activity of the subject, number of sessions and the time per session of the data. Some of these parameters we already explored in section 2 because can lead to different approaches namely on-the-person or off-the-person approach.

To develop our system, it is of the highest interest to ramp up a database with all the parameters referred before in order to fit the real-life application.

5.3 Long-term ECG Database

Long-term ECG databases record the hearts rhythm over a considerable period of time, such 6 months. Comparatively to the 24-hour ECG recorder this is a less sophisticated recorder but allows measuring the rhythm during different phases. An alternative consist in an invasive procedure for implant a loop recorder, like pacemaker (able for recording periods of 3 or more years). In Tab. 5.1 we present some of the long-term databases found in the literature.

E-HOL 24h Holter is documented as E-HOL-03-0202-003 by the University of Rochester[1]. Consists of a study of 202 healthy subjects. Furthermore, the individual records were acquired using three leads, where exist an initial resting supine period for 20 minutes duration before starting the ambulatory recording with a sampling frequency of 200Hz.

The Long-Term ST Database is available in physionet and is described by (Jager et al., 2003). Contains 86 ECG recordings of 80 human subjects, chosen to exhibit a variety of events of ST segment changes such as ischemic and non-ischemic. The database was created to develop and evaluate the performance of algorithms to detect mechanisms and dynamics of myocardial ischemia. The individual recordings were acquired with two/three leads, digitized at 250Hz and are between 21 and 24 hours of duration.

5.4 State-Of-The-Art: Template Update

In real life scenarios, it is faced several challenges, namely the variations of the input biometric data, due to several parameters already discussed in Chapter 4, such as ageing of the biometric traits, and some variations due to the interaction individual-sensor. This will contribute to a large intra-class variability witch lead to a poor representatives templates of biometric data to be recognized, resulting in poor recognition performances. Continuous verification requires a continuous update of the template to maintain the performance results over time.

This step consists in identity labels and can be done by supervised learning (i.e by human experts) or by semi-supervised (i.e by the system). The significant difference between these methods is the technique followed for the data labelling.

In the literature, it is difficult to find a rule for template update (it is not yet a very studied research field) but some authors use the acceptance threshold equal to zero false acceptance rate point because assure the selection of only genuine samples for template update (Komeili et al., 2018). Nevertheless, it is possible to identify some common mechanisms that may vary depending on different factors:

- Mechanisms such as the choice of the update criteria (threshold, graph-based);
- The periodicity of the template update (online or offline);
- The template selection mechanism;
- Template update working mode system (supervised or semi-supervised).

All techniques that will be explained in this chapter are summed up in Fig. 5.1.

Table 5.1: Comparison between literature long-term databases based on 8 parameters.

Database and Origin	**Publicly**	**No. of Subject**	**Sampling Rate (Hz)**	**Leads/ Electrodes**	**Health Conditions**	**Activity/ Posture**	**Duration**	**Session Time**
BIDMC (Baim et al., 1986)	y	15	250	chest	heart failure	ambulatory ECG	-	20h
Long-Term ST (Jager et al., 2003)	y	80	250	2/3 lead	-	ambulatory ECG	86 records	21h to 24h
ECGDMMLD (Johannesen et al., 2016)	y	22	500	12-lead	health	supine position	24h	10s each sample
ECGRDVQ (Vicente et al., 2015)	y	22	500	12 leads	healthy	supine position	7 days	10s each sample
ECG-ID (Yadav et al., 2017)	y	90	500	single-lead	-	-	6 month (+- 20 sessions)	20 sec
MIT-BIH Long Term (Laguna et al., 1997)	y	7	250	chest	-	-	14h to 22h	5 min
Long-Term AF (Petrutiu et al., 2007)	y	44	128	-	-	-	24-25h	1 min
PAF (Moody et al., 2001)	y	100	128	2-channels	-	-	100 records	each record (2 of 30 min and 2 of 5 min)
Apneia-ECG (Penzel et al., 2000)	n	70	100	chest	-	sleep	27 records	401 to 578 min
CYBHi (da Silva et al., 2014)	y	128	1000	palms + finger 2 leads	none	-	3 month	5 min
UofTDB (Wahabi et al., 2014)	y	1020	200	finger(1/2)	none	-	6 month	2 to 5 min
Lviv (Wieclaw et al., 2017)	y	18	277	3 finger leads	none	-	15 days	7 records (10s)
SHAREE (Melillo et al., 2015)	y	139	128	chest	hypertensive	-	12 months	5 min
E-HOL 24h (University of Rochester[1])	y	202	200	chest	none	ambulatory recording	-	24h session

[1] http://thew-project.org/Database/E-HOL-03-0202-003.html

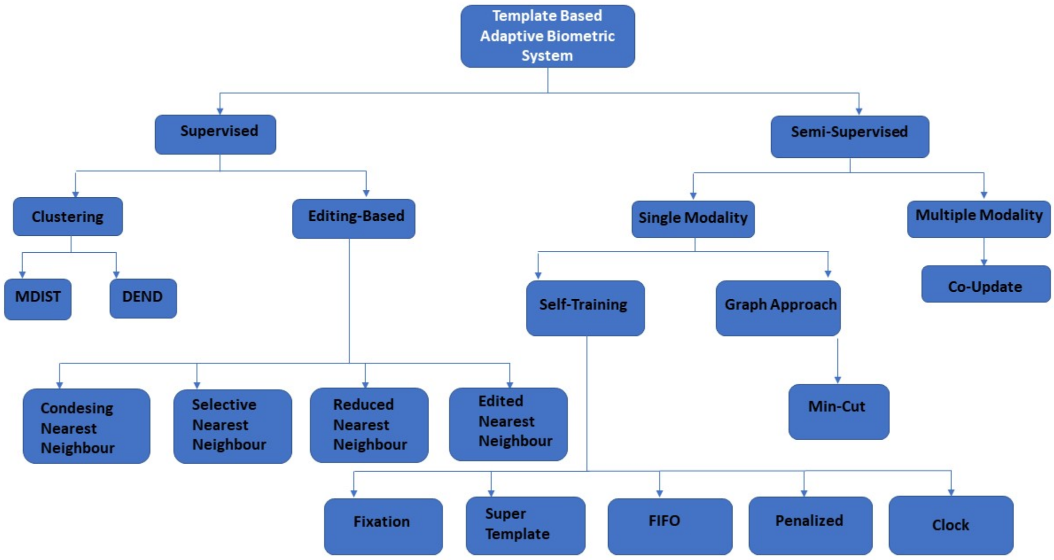


Figure 5.1: Dendrogram regarding taxonomy of template update techniques based on (Rattani et al., 2010).

5.5 Supervised and Semi-supervised Learning Methods

5.5.1 Supervised Methods

Supervised methods are offline methods which use a feature vector and a class label (target). The label attribution is given by the supervisor. This actor needs to have knowledge of the correct intra-class variation. Here we guarantee no impostor data has been used for the template update (Akhtar et al., 2014).

5.5.2 Semi-supervised Methods

One derivative from machine learning approaches is semi-supervised methods, that merge labelled (in biometrics represents the initial training samples) and unlabelled (correspond to the samples available during system operation) data for improving the system performance (Zhu, 2008).

The system updates the template iteratively using only high confident input data (matching score with the current template is above a certain threshold) and attributes labels using the knowledge learned by the model (in order to avoid impostors). Since they update themselves in an automatic way, they are also called "self-update" methods (Marcialis et al., 2008).

5.6 Clustering Methods

According to Lumini and Nanni (2006) these methods are based on applying standard clustering algorithms:

• *MDIST* aims to search for the templates that minimize the intra distance among all the samples in the database (i.e the most similar). This algorithm comes from the distance-based clustering and has as selection criteria the very close samples;

• *DEND* aims to search for the templates that exhibit large intra-class variations resort to the dendrogram (i.e the most different). This algorithm is also coming from the hierarchical clustering and contrary to the MDIST, selects the farthest samples.

These methods were deeply studied for fingerprint recognition by Uludag et al. (2004) and combine these two methods with:

• *BATCH-UPDATE:* Similar to FIFO technique, where discards the previous template set and take into account only the latest acquired data.

• *AUGMENT-UPDATE:* Considers the freshly acquired data as well as the actual template to perform template selection.

They obtained best results in *AUGMENT* in *MDIST* methodology. With this technique they can retain long-term trends by allowing previous selected templates compete for reselection.

Uludag et al. (2004) were unable to find the ideal number of templates per individual with the *AUGMENT-UPDATE* which can lead to unsustainable computational cost (i.e the number of templates is always growing). In this method, the selection of the best templates by the current collected template of the individual is crucial for a real-life system.

In Li et al. (2008) a study on biometric template selection (improve the system's accuracy efficiently if enough samples are captured in identification) and update is performed. Here it is performed a similar work of Uludag but emphasis is given to the number of templates that should be selected (contrary to Uludag et al. (2004)) and put the condition of choosing K templates that need to be less than N, the total number of templates. To choose the value of K they propose a method called model maximized score model, that is equivalent to $K \leq [\frac{N}{2}]$ if N is large enough.

5.7 Editing-based Methods

Hutchison et al. (2008) and Freni (2010) proposed the use of editing methods because these give focus on the whole collected training set T. They generate a subset E ($E \in T$) that maintain the same classification performance of T. The best subsets were obtained by review the structure of the data (needs to be done to each subject). These kind of method are independent of the number of templates.

Freni (2010) proposed some algorithm methods based on K-Nearest Neighbour classifier (KNN) a non-parametric classifier with high computational complexity. All the algorithm must be representative of T and can be roughly described as *incremental* when the E starts empty and grows, or *decremental* when E starts equal as T and in each iteration some instances are deleted until some criteria are reached, such as:

• *Condensed Nearest Neighbour* is a *incremental* algorithm that aims to find E such that the closest instance $y \in T$ relative to the instance $x \in E$ exhibits the same class identifier of x. The maximum classification accuracy on T is reached (Abboud and Jassim, 2012).

• *Selective Nearest Neighbour* is also a *incremental* algorithm that aims to build E as small as possible. The instances that are as much close to the owner class than any other instance of different class are added to E. That not guarantee the 100% classification accuracy on T (Ritter et al., 1975).

• *Reduced Nearest Neighbour* is a *decremental* algorithm. Starts by iteratively eliminating instance from E (since that not imply misclassification on T). This algorithm stops when no other instance can be removed from E (Gates, 1972).

• *Edited Nearest Neighbour* is a *decremental* algorithm. When an instance did not match with the majority of their k nearest neighbours it is removed from E (Wilson, 1972).

This support the idea of the biometric menagerie (already described in chapter 4), that different subjects can provide a different level of difficulty in recognition and can be grouped in order to enhance the system performance. According to Rattani et al. (2010) editing approaches could help to decide a variable number of templates for each individual. They propose detecting individually the correct number of templates needed for good classification accuracy, and for achieving that, they label easy or difficult the subject.

5.8 Semi-Supervised methods

5.8.1 Self-Update methods

These techniques are based on semi-supervised learning, where they self-update the templates based on the intra-class variation of the new data, comparing to the template stored in the database (Zhu, 2008) (Akhtar et al., 2014). Only select the highly confident data, following the heuristic search (Threshold condition). These kinds of approaches also need a stop criteria, otherwise, the algorithm will never stop.

5.8.1.1 Offline methods

Offline methods are adaptive methods, in which the needed data is collected over a period of time. The samples collected are iteratively used to update the enrolled templates. After each interaction, the value of the threshold is re-estimated. The methods based on supervised learning are usually offline methods while semi-supervised methods can be either online or offline (Freni et al., 2008) (Marcialis et al., 2012) (Akhtar et al., 2014).

5.8.1.2 Online methods

Contrary to offline methods, the templates are updated to the rhythm that new data is collected. Here, the match between the previous data and this new one is compared and if the score is bigger than the threshold, the data will be used for update the templates (Xudong Jiang and Ser, 2002) (Freni et al., 2008) (Marcialis et al., 2012) (Akhtar et al., 2014). These methods are dependent on the order of the input sequence arrives. The performance of the update algorithm will behave differently if the same input data set is observed but the order is different.

5.8.1.3 Techniques description

FIFO approach: The most common strategy is the first-in-first-out, where a threshold is defined and when our sample is above that threshold, the database is updated (discarding the previous sample) (Coutinho, 2011) (Niinuma and K. Jain, 2010) (Komeili et al., 2018). This technique is no time-consuming. Another similar approach was proposed by Lourenço et al. (2011) where an individual is accepted if the Euclidean distance between the template test and the template was less than a threshold.

Labati et al. (2014) proposed creation of a *"super template"* X composed by N templates x, that means that the new genuine date is always fused to a common single template called "Super Template" containing all the data.

The super template is updated online during the execution of continuous verification. The update condition arises of a previous threshold value that avoids impostors, and the authors recommended the use of zero False Match Rate or EER.

The best results emerged when they tried to maximize the intraclass variability covered by the templates belonging to the super template.

Chun (2016) proposed a *Penalized template update*, a technique based on the mean of the past ECGs and the actual ECG. They use a cost function, where they penalized maximum likelihood estimator to minimize the cost function by penalizing exponentially the previous ECG (attributing weights to each ECG sample). Basically, the more the ECG is updated, the less impact the older ECG will have.

Guerra-Casanova et al. (2011) proposed a technique based on the *fixation* of samples while the remaining are updated. In this work, they save three initial samples and perform several tests.

The best results were obtained when they update alternatively one of the three samples of the template comparatively to when one or two samples of the template remain without alteration.

Scheidat et al. (2007) proposed a *clock method*, a technique similar to FIFO replacement strategy. Works by having a circular list of samples and each element in the list stores a template (contains at least the sample number and the reference value, which is 0 or 1). All templates start with reference value 0 and when is accessed by the system (the best match) that value is set to 1, iteratively. When complete a cycle the systems search the reference with value 0 and substitute by the new one. In the end, all the templates are set to 0, to start a new cycle.

5.8.1.4 Problems with Self-update Methods

For updating, the data that will be accepted if the matching score is above the selected threshold, witch lead to a dependency of these methods to a threshold value (Rattani et al., 2010).

The threshold is estimated using initially enrolled templates. This leads to some problems:

- These methods can miss some important intra-class variations because they use only the patterns similar to the templates stored;
- Online methods are dependent on the order of the sequence of input data;
- Vulnerable to large intra-class variations;

Table 5.2: Comparison between the graph-based methods and their respective functions

Method	Source	Loss	Regularizer
Min Cut	(Blum and Chawla, 2001)	$\infty\sum_{i\in L}(y_i - y_{iL})^2$	$\frac{1}{2}\sum_{ij} w_{ij}(y_i - y_j)^2$
Gaussian Random Fields and harmonic Function	(Zhu et al., 2003)	$\infty\sum_{i\in L}(f_i - y_i)^2$	$f^T\Delta f$
Local and Global Consistency	(Zha et al., 2009)	$\infty\sum_{i=1}^{n}(f_i - y_i)^2$	$f^T D^{-\frac{1}{2}}\Delta D^{\frac{1}{2}} f$
Tikhonov Regularization	(Belkin et al., 2004)	$\frac{1}{K}\sum_{i}(f_i - y_i)^2$	$\gamma f^T S f$
Manifold Regularization	(Sindhwani and Niyogi, 2005)	$\frac{1}{l}\sum_{i=1}^{l} V(x_i, Y_i, f)$	$\gamma_A\|f\|_k^2 + \gamma_I\|f\|_I^2$
Graph Kernel from the Spectrum of Laplacian	(Chapelle et al., 2006)	$\min \frac{1}{2} w^T W$	$\exp(-\frac{\sigma}{2}\lambda)$
Spectral Graph Transducer	(Zhu et al., 2003)	$\min c(f-\gamma)^T C(f-\gamma)$	$f^T L f$
Local Learning Regularization	(Kokkinos and Margaritis, 2017)	$\min \frac{1}{k}\sum_{i=1}^{k}(y_i - f_k(x_i))^2$	$\frac{\gamma}{k}\|f_k\|^2$

• Since the algorithm normally looks for the minimal cost (high score), it can be stuck in local maximum;

• Always use highly confident data for updating.

5.8.2 Graph-based Methods

5.8.2.1 Introduction

These methods are commonly define as a graph where the nodes are labelled (system knows who is the subject) and unlabelled (the system do not know who is the subject) data and the edges are the similarity between those samples. The edges can have different weights (Zhu, 2008) (Rattani et al., 2008) (Rattani et al., 2010).

To be seen as graph-based semi-supervised methods the method must estimate a function f on the graph and satisfy two rules (Zhu, 2008) (Rattani et al., 2008) (Rattani et al., 2010):

1) Must be near to the given labels Y on the labelled nodes;

2) Must have two terms to turn the graph smooth:

• First term is a loss function;

• Second term is a regularizer.

Several similar graph-based methods exist and the only difference is the choice of these two terms (as can be seen in Fig. 5.2) (Zhu, 2008).

Since the problem of template update is an optimization labelling (tries to find the maximum intra-class variation), it is possible to modulate into a graph approach, where the graph nodes represent samples (labelled and unlabelled) and the edges represent the weight (score) between a pair of samples.

Among all, few applications of template update for biometrics are found in the literature. The most studied method is the min-cut approach.

5.8.2.2 Graph Min-cut

Blum and Chawla (2001) proposed a graph min-cut problem (also known as st-cut), where given a set L of labelled data (L+ denotes a positive example, and L- denotes a negative example) and U represents unlabelled data. Aims to find a minimum set of edges whose elimination stops the flow from the L+ (source) to the L- (sink).

Rattani et al. (2008) exploited an updating mechanism that can adapt the user model to all the possible intra-class variations available in the unlabelled data in facial recognition. The classification of unlabelled input data is done by the division of the graph. These methods consist of techniques of label propagation, namely using the minimum weight of edges. The proposed technique may capture large intra-class variations without increasing the probability of impostor introduction (important to refer that in this paper it is exploit authentication problems).

In the same publication, the authors also proposed a based graph min-cut approach for biometric template update (offline). They combine the stored and the input data in a identically graph-structure based on pairwise proximity where nodes represent combined data and edges represent the resemblance information among two nodes. Also described as an uninformed search algorithm, that searches all the possible combinations, where the start node is the template and the goal node is large intra-class variation.

5.8.3 Co-update Methods

The "self-update" methods have some handicaps as seen in the previous section. Template co-update was inspired by semi-supervised learning method called co-training to increase the performance of the template update of multimodal biometric problems (biometric systems that combine more than one biometric trait). Basically, two matches are trained with the same template of the subject. One matcher operates at high confidence (considers unlabelled samples as highly truthfully), and this will help the other to identify difficult patterns (Roli et al., 2007).

Co-update is used to capture large intra-class variations, thanks to the complementary biometric.

5.9 Evaluation Metrics

Giot et al. (2012) proposed two evaluation metrics for template update performance:

• Impostor Update Selection Rate (IUSR) coincide the proportion of impostor's samples involved in the update process among all the tested impostor's sample, where U_i represents the number of impostor's samples and N_i all tested impostor samples as can be seen in Fig. 5.1

• Genuine Update Miss Rate (GUMR) corresponds to the proportion of genuine samples not involved in the update process among all the tested genuine samples where U_g represents the number of genuine samples selected in the updating process and N_g all the genuine samples tested as can be seen in Fig. 5.2.

The authors stated that the best template update systems stay when the samples are noisy (our case since it is expected to use off-the-person signals) the GUMR should be closer to zero and should miss some noisy genuine samples. They also stated that IUSR should be zero.

$$IUSR = \frac{U_i}{N_i} \tag{5.1}$$

$$GUMR = \frac{N_g - U_g}{N_g} \tag{5.2}$$

5.10 Model Update

Since template update is for systems based on the matching with distances as seen before, it is required other technique for the others methods. When we are dealing with classifiers it is needed *model update*, and consist in re-training the model used for classification, for example when it is used SVM (this model must be trained offline, otherwise will have an unacceptably high computational cost), kNN (it is needs to add the new templates to the classifier dataset that can lead to memory issues) or ANN (it is needs to do backpropagation with the new templates in order to update the model).

5.11 Conclusion

When we are developing an ECG biometric system we need to choose carefully the database since an ECG signal collection ensures the quality of the performance assessment of a biometric system. Some factors of the database can have influence, such as the number of subjects, sampling rate, number of leads/electrodes, health conditions, activity/posture, duration and session time.

With this in mind, a public database with the longest recorders of the subject ECG will be used during the dissertation, in order to test with reliability the template update advantages, namely E-HOL 24h.

Besides in the literature, they obtain good results with medical signals, the performance decay over time. After the discussion on the previous chapters, it is now understandable why ECG biometric systems fail, namely due to ECG variability over time, the acquisition is not perfect and leads to signal noise. One possible solution for the problem can be template update, where it is continuously updated the user template. These techniques are quite intuitive and consist in making the biometric recognition systems adapt over time.

Despite the huge effort in the topic we can conclude that the basic operations regarding template update are addition, deletion, and compress or merge. Besides, there is the opportunity to exploit these methods in ECG biometric systems, since they were already exploited in other biometric traits.

Chapter 6

Variability Study

6.1 Introduction

After the review of the literature such as in Ye et al. (2010), where they study if the ECG signal was good for biometric purposes. They test the system in three different databases and conclude that in a long-term database (Long-Term ST database, described in Chapter 5) the algorithmic performance decomposed over time.

Guerra-Casanova et al. (2011) perform a study on gesture recognition with template criteria/update. They verify that a template updating strategy is mandatory. They present an FRR of 70% of error when a long time has passed and no updating process has been carried out. With this in mind, we consider template update as one possible solution for long-term biometric approaches.

To fully and objectively evaluate the effects of ECG variability on the performance of biometric algorithms we conduct a study where we tested four literature methods: the method described by Plataniotis et al. (2006), Tawfik et al. (2010), Belgacem (2012), Eduardo et al. (2017) and tested in a house method. These state-of-the-art methods were implemented and tested at different points in time on the available databases.

In Plataniotis et al. (2006) an ECG biometric recognition method was proposed using a non-fiducial approach using PTB database with windows of 10 seconds, at the pre-processing phase a band-pass filter (0.5 to 40 Hz) is used. Feature extraction is performed with autocorrelation and dimensionality reduction by using a discrete cosine transform (DCT)(the fifteen most relevant features were selected). Euclidean distance is used for classification.

Tawfik et al. (2010) used in the pre-processing phase a band-pass filter (1 to 40 Hz). They identify the QRS complex and cut it from the signal with a window of 0.35 seconds (because it is the part most stable from the ECG), perform the mean of those QRS signals and extract features using DCT technique (the thirty most relevant features were selected) . A multilayer perceptron (MLP) is used for classification.

Belgacem (2012) also used a pre-processing with a pass-band-filter (1 to 40 Hz). They identify the QRS complex, cut them from the signal, compute the mean and do the feature extraction resorting to Discrete Wavelet Transform (DWT). Since DWT decomposes the signal in smaller

(a) Saturated Signals. (b) Unusable Signals.

Figure 6.1: Inconclusive Signals.

parts (in our case in 4 levels) it is used for characterization of the subject only the most relevant (that contain more expressive features). A Random Forest is used for classification.

Eduardo et al. (2017) used a Finite Impulse Response (5 to 20 Hz) filter with 150 order for pre-processing. They detect and cut with a fixed length of [-200, 400] ms around R peak. Detect and remove the outliers with DMEAN ($\alpha = 0.5$ and $\beta = 1.5$ and with Euclidean distance). For classification they resort to Knn with k number of neighbours equal to 3 and with cosine distance.

(BOOK CHAPTER REF) it was proposed an Convolutional Neural Network (cNN) for biometric identification based on non-intrusive ECG acquisitions. (EXPLICAR A CNN E COLOCAR UMA IMAGEM, referir que as condições de teste aqui são a cada 5 segundos o que contém menos informação e por isso menos robusto, e como temos melhores resultados é de realçar este facto)

During this master thesis,in order to make easier the communication, we will reference as Plataniotis method the method proposed by Plataniotis et al. (2006), Tawfik method the one proposed by Tawfik et al. (2010), Belgacem method the one proposed by Belgacem (2012), Autoencoder method the one proposed by Eduardo et al. (2017) and as cNN method the one proposed by (BOOK CHAPTER REF).

6.2 Experimental settings

We use the signals from E-HOL 24h Holter as explained in Chapter 5. Each subject contain 3 signal leads but is not coherent for all the subjects (that means that the lead 1 in one subject can be lead 0 in other), what forced us to see individually what lead is more similar to the traditional lead 1 that contain a better ECG signal for biometric purpose. As soon we go deep in this database and besides the fact that each subject has 3 leads, we realise that some subject signals are not fit for the job. It is possible to see that some signals are saturated (see Fig 6.1a), or with to much noise (see Fig 6.1b). Finally we decide to remove 13 subjects signals from the origin database what lead us with 188 patients (it was removed subject 1043, 9003, 9005, 9020, 9021, 9022, 9025, 9046, 9061, 9064, 9071, 9082 and 9105).

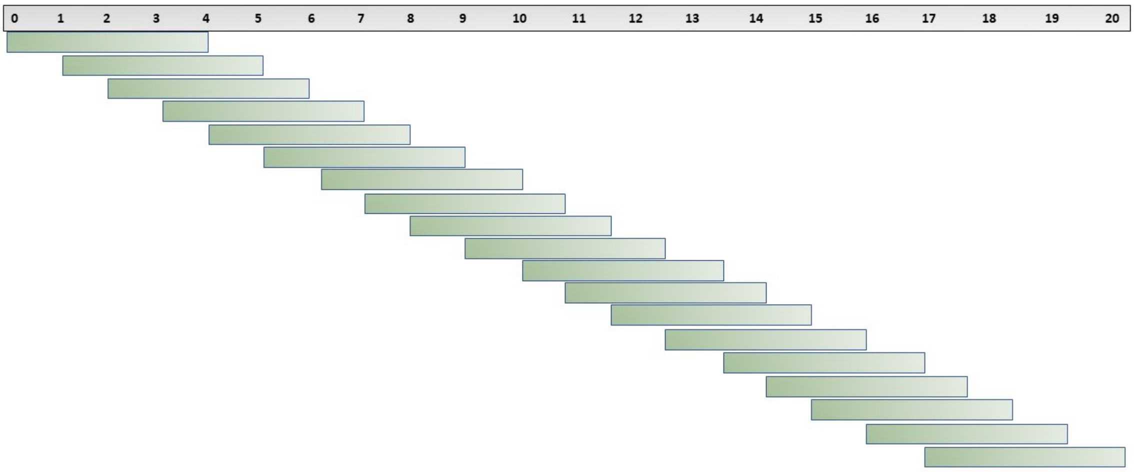


Figure 6.2: Signal Overlap: Technique used for increase the training data while ensuring realistic settings. The signal was cut in small segments of 5 seconds.

Besides, we have replicate the work described in Labati et al. (2013) for evaluate the permanence of ECG over time. It is used 5 minutes for the training and 5 minutes for testing. In this work, the authors use the 3 leads of each subject and perform a noise reduction (notch filter 50 Hz) and a High-Pass Butterworth filter (0.5 Hz). Then, identify R peak, cut the QRS complex (window of cutting [-70, 50] ms) and do the mean. After, perform the matching and fusion score (of each lead). It is important to refer that this author perform a study of ECG signal in biometric but for authentication.

6.3 Signal Pre-processing

We start by dividing the dataset in train set and in the test set. During the training phase, the last 30 seconds of data (see Fig. 6.4a) were used (close to the real-life application) of the first 60 minutes (more stable signal comparing to first 20 minutes according to University of Rochester[1]) of the signal with overlap (see Fig. 6.2) with the aim of increase the training data (in Fig. 6.4b it is possible to see a sample of 5 seconds). For the testing, a study was performed over time. Starts immediately after the training and after some time it was tested again as can be seen in Fig. 6.3. We use 15 minutes of data but we divide this 15 minutes in intervals of 30 seconds.

After this first step we eliminate the samples that aren't complete (see Fig. 6.4c and Fig. 6.4d)

Continuing, in the Plataniotis method it is done a Butterworth filter (order 3) with a band-pass of 0.5 Hz to 40 Hz (see Fig. 6.5a), for cNN method we use raw data and the other three methods with a cut-off of 1 Hz to 40 Hz (see Fig. 6.5b). Here we felt the need of change a the

[1] http://thew-project.org/Database/E-HOL-03-0202-003.html

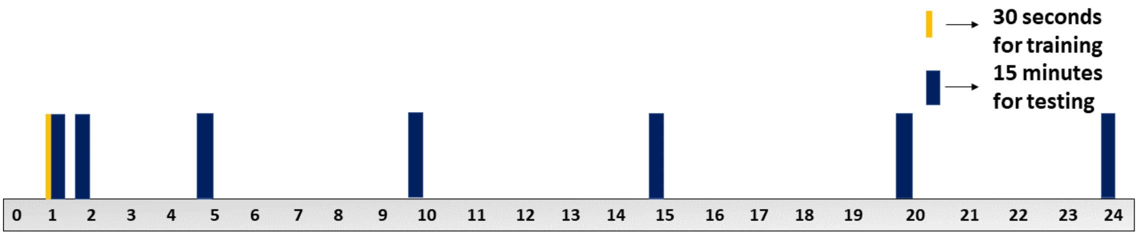


Figure 6.3: The blue and red rectangles represents the intervals of the signal that were used for training and classification.

cut-off frequency of Autoencoder method because the algorithm was losing important frequency for biometric purpose.

Regarding the normalization we present the results of Plataniotis, Tawfik and Belgacem method resorting to the function 6.1. On the other hand to satisfy the code requirements of the Autoencoder method we resort to a different normalization as can be seen in function 6.2. The results are shown in Fig 6.6.

$$Input[g] = \frac{input[g] - \overline{input[g]}}{\sigma input[g]} \tag{6.1}$$

$$Input[g] = \left(\frac{input[g] - \min input[g]}{\max input[g] - \min input[g]} \right) \times 2 - 1 \tag{6.2}$$

6.4 Feature engineering

Continuing, for feature extraction in Plataniotis method it is used autocorrelation and for feature selection it is used Discrete Cosine Transform (DCT), where (as can be seen in Fig. 6.7) was selected the 15 strongest features.

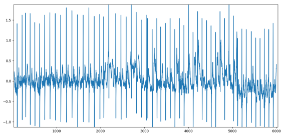
In Tawfik method it is used DCT for extract the 25 strongest (see Fig. 6.8).

As Belgacem, we use DWT to extract features, but we divide the signal in 4 levels because our Sampling Rate comparing to the data used in their experiment e lower, and in this way we capture the most relevant frequency for biometric purpose (we use cd4, cd3, cd2, cd1) (see Fig 6.9).

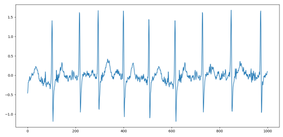
Before perform the extraction of features in Auto-encoder method we do outlier removal with $\alpha = 1.2$ and $\beta = 1.5$ comparatively to the state-of-the-art values for $\alpha = 0.5$ and $\beta = 1.5$, because it is more suitable for our data. After we perform Auto-encoder (batch size of 256, 100 epochs with 3 Dense layers, with an optimizer $ADAM = 0.01$) where we reduce from 120 features to 20 (see Fig 6.10) with a loss in order of 10e-4.

6.5 Classification

This step it is done with kNN for Plataniotis ($kneighbours = 1$) and Auto-encoder ($kneighbours = 1$ contrary to the $kneighbours = 3$ present in the state-of-the-art), with MLP for Tawfik (20 hidden



(a) Input Signal.



(b) Training signal with 5 seconds (overlap).

Name	Type	Size	Value
x_test_10h	list	5580	[Numpy array, Numpy array, N…
x_test_15h	list	5520	[Numpy array, Numpy array, N…
x_test_1h	list	5610	[Numpy array, Numpy array, N…
x_test_20h	list	5481	[Numpy array, Numpy array, N…
x_test_24h	list	2358	[Numpy array, Numpy array, N…
x_test_2h	list	5610	[Numpy array, Numpy array, N…
x_test_5h	list	5580	[Numpy array, Numpy array, N…
x_train	list	4862	[Numpy array, Numpy array, N…

(c) Size of the samples.

Name	Type	Size	Value
x_test_10h	list	5580	[Numpy array, Numpy array, N…
x_test_15h	list	5520	[Numpy array, Numpy array, N…
x_test_1h	list	5610	[Numpy array, Numpy array, N…
x_test_20h	list	5480	[Numpy array, Numpy array, N…
x_test_24h	list	2328	[Numpy array, Numpy array, N…
x_test_2h	list	5610	[Numpy array, Numpy array, N…
x_test_5h	list	5580	[Numpy array, Numpy array, N…
x_train	list	4862	[Numpy array, Numpy array, N…

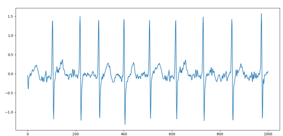
(d) Size of the complete samples.

Figure 6.4: Common step in ECG signal pre-processing.

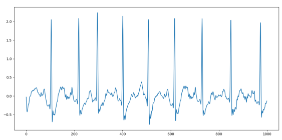
layers are used) and with Random Forest to Belgacem (20 trees are used). A study was conducted by synchronize the start recording time of all the subjects, resort to the Plataniotis method to evaluate the behaviour of such approach.

6.6 Results

Comparing our results with the ones shown by Labati et al. (2013) we conclude that the ECG signal isn't fully permanent over 24 h. The permanent showed up related to the time and keep relative good results in the first two hours (see Fig. 6.11). The results obtain were first obtain per batch, that mean for each of the 30 seconds of the 15 minutes, where we conduct a weight mean according to size of the batch (there was signals shortest that others). It is possible to see the results relative to the Plataniotis, Belgacem , Tawfik and Auto-encoder in Fig. 6.12. The performance is low compared with the results obtained by these four studies, at least for the immediate test (first test after train) because it was expected to obtain results above 97% of accuracy (can be explained by the small differences of implementation that can exist to the articles because of missing information and by the data), it is visible that the performance decays over time significantly. Notwithstanding, from these results, we can conclude that the variability changes considerably over the day.

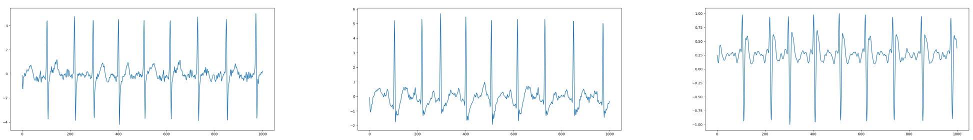


(a) Plataniotis method after Butterworth filter.



(b) Tawfik, Belgacem and Autoencoder method after Butterworth result.

Figure 6.5: Signals filtering resorting to Butterworth filter.

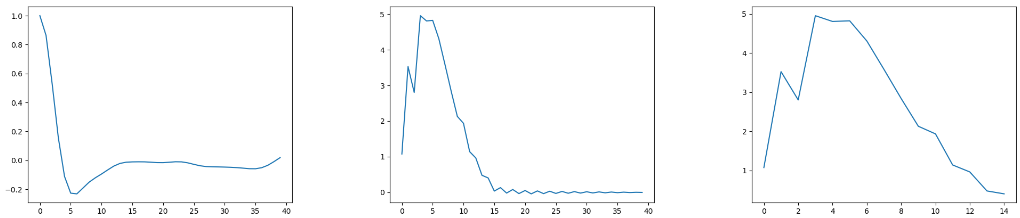


(a) Plataniotis method after normalization.

(b) Belgacem and Tawfik method after normalization.

(c) Autoencoder method after normalization.

Figure 6.6: Signals normalization.



(a) Plataniotis method after autocorrelation.

(b) Plataniotis method after DCT.

(c) Plataniotis method after choosing the strongest features.

Figure 6.7: Signals progression according to Plataniotis method

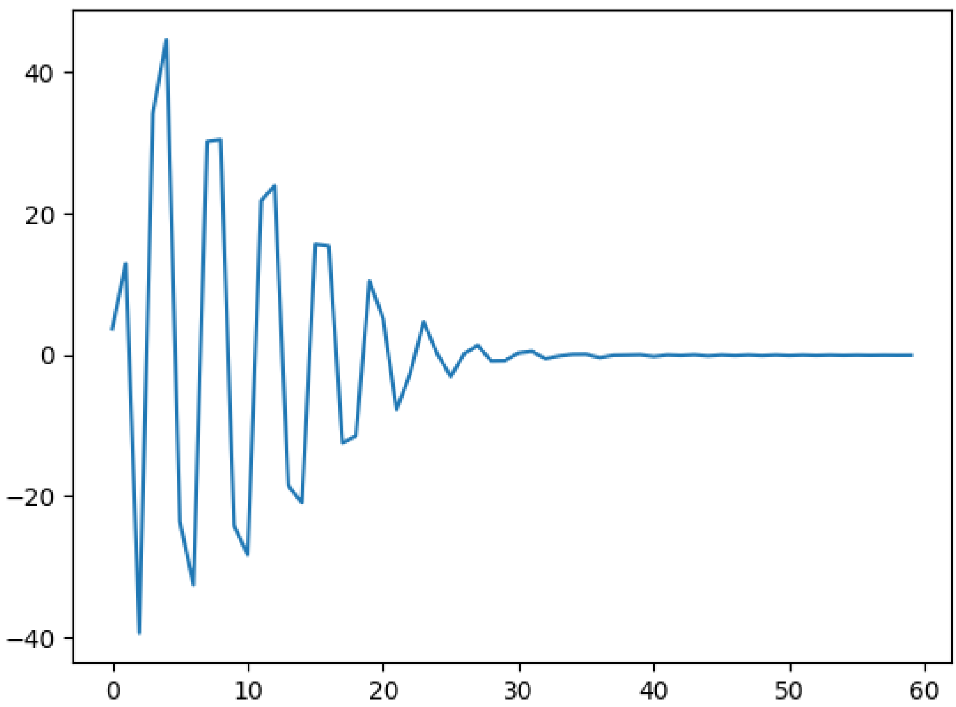


Figure 6.8: Tawfik method feature extraction.

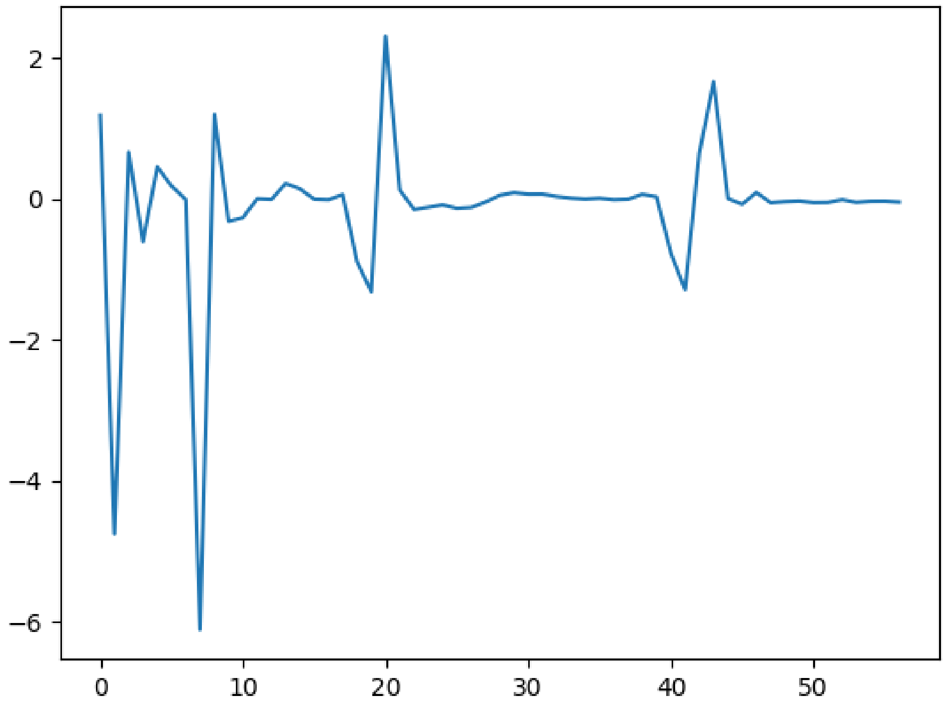


Figure 6.9: Belgacem method feature extraction (it is possible to identify the 4 levels).

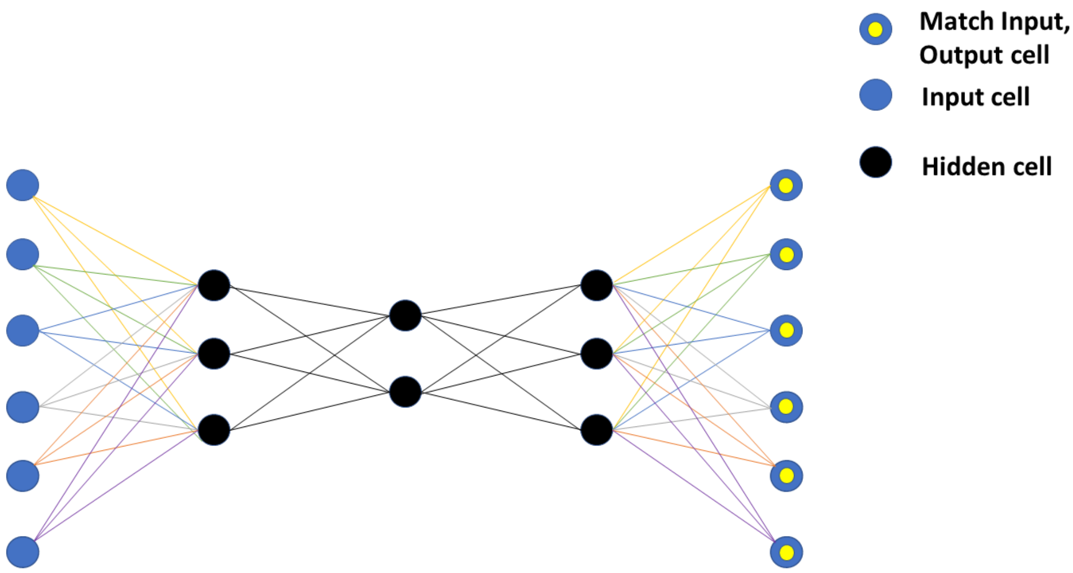


Figure 6.10: Structure Auto-encoder.

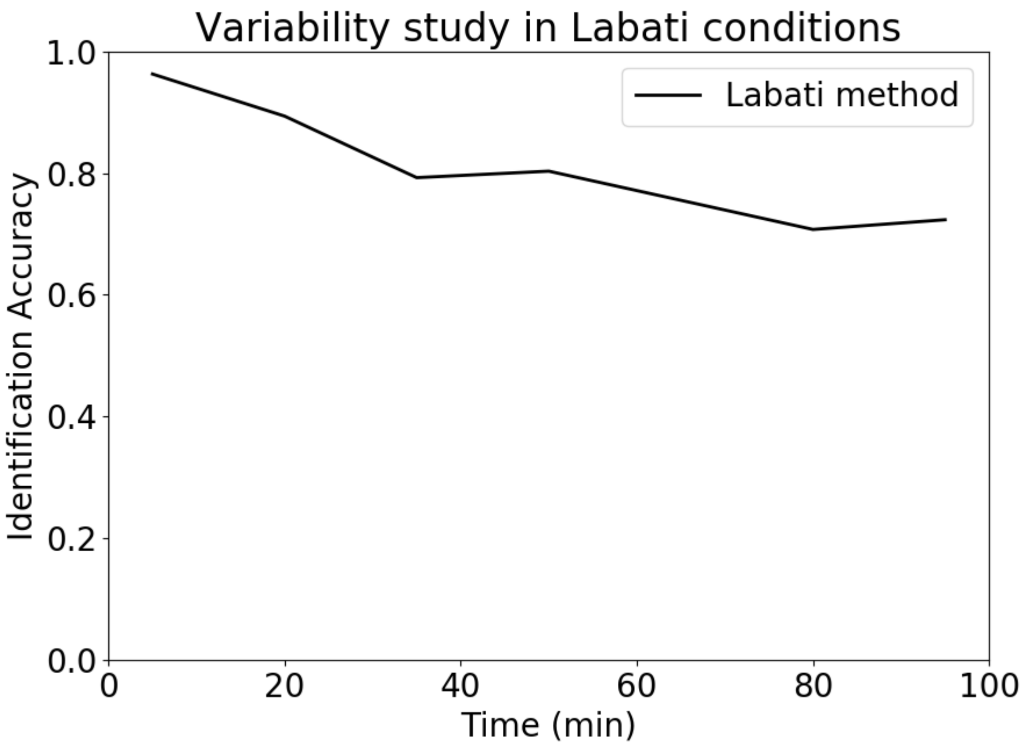


Figure 6.11: Labati Results for the first two hous.

The results of synchronize starting point (see A), reveals consistent with the results shown by the random starting point. Can be denoted that the performance suffers a considerable decay past 15 hours and after starts to increase (low increase but significant). The majority of the subjects start the recording between 8-12 a.m, and past $15H$ correspond to the interval of 11 p.m and 3 a.m. Extrapolating the data for a normal routine of an healthy individual, it is most likely that at that period the subject is sleeping. It is possible to conclude that our ECG behaves very different from the sleeping state, where it appears to reach the maximum of variability during the NREM sleep (probably during the 3-stage "deep-sleep" where the EEG slows dramatically their activity).

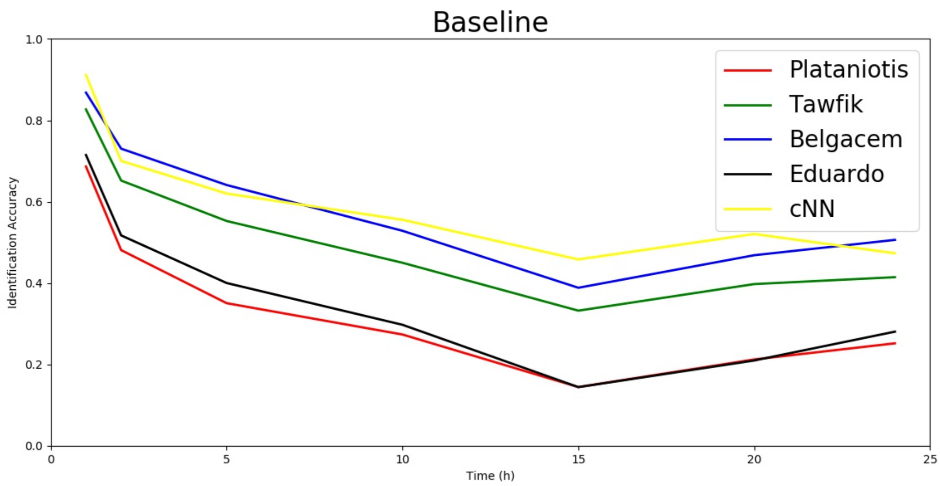


Figure 6.12: Variability Results, corresponding to red Plataniotis, blue Belgacem, in green Tawfik, in black Auto-encoder and in yellow cNN.

Chapter 7

Template Update Experiments

7.1 Introduction

After the study of variability and permanence of ECG signal and as we seek for a real-life applicability and real-life performance, the results obtain are not good enough. So in this Chapter we intend exploit several techniques of template update in order to improve the results obtain. It is expected that with this techniques the model will learn the variability and increase the performance.

7.2 FIFO

7.2.1 Threshold

The first technique that we implement was FIFO. This technique needs a rule for allowing the update occur. As so, we use the training information (enrolment data) for search the ideal threshold. We use 3/4 of the size of the training samples for train a model and use 1/4 of the size or testing and we plot the error to see with what threshold we obtain a better ratio of TP and FP (an example is shown in Fig 7.1).

With the plot, we where able to choose several thresholds in order to keep the one that gave us better results (near of the image yellow area).

7.3 Fixation

Considering the results obtain in FIFO we choose the threshold with better results and use in the Fixation technique. Here, we developed a generic technique where it is fixated $n/4$ samples of the size of the total samples of the individual, where $n \in [1,2,3]$ and the rest of the samples are updated. Similar to this approach we use one more approach where we choose the samples that are fixated though a matrix of cross-correlation and test with $n/4$ most and less similar.

Still in this method, we implement in a more controlled environment, another approach. It is fixated $j \times n + n$ samples where $n \in [1,2,3]$ and j represent what test are we conduct, for example in our case we have seven moments of tests, therefore, $j \in [0,6]$. If a system based in this method

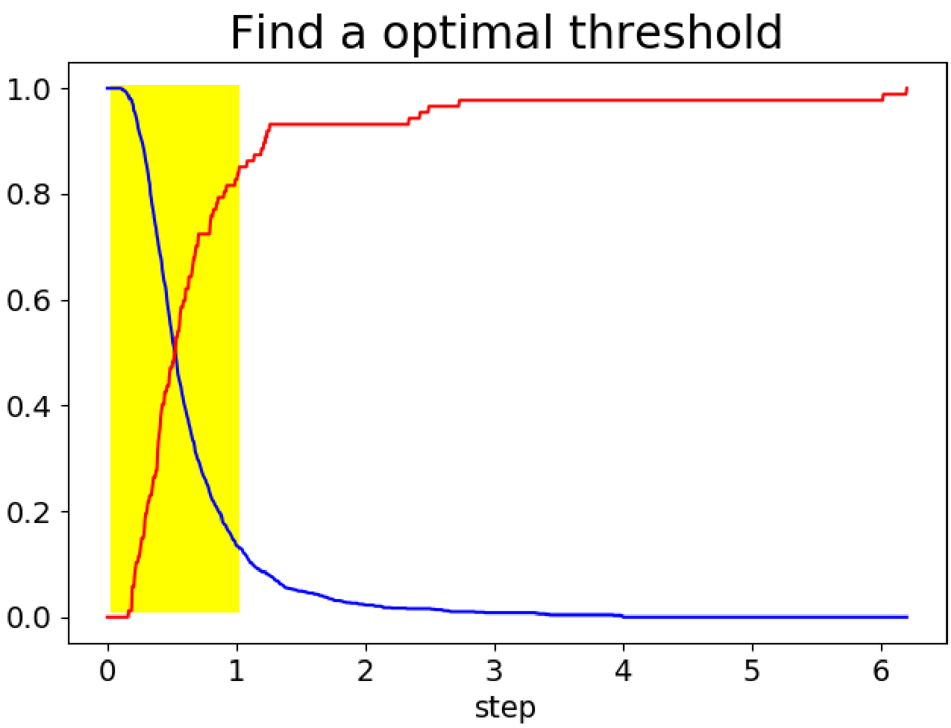


Figure 7.1: Search for the ideal threshold.

is intended to a real-life application, will need a few changes in the implementation. This effort was conducted for the model learn better the variability over time.

7.4 Adaptive Clock

After analyse the approach of Scheidat et al. (2007), we decide to implement a similar algorithm. In Eq. 7.1 is described our approach. The distance between the actual sample and the ones in the database are computed. The oldness of a sample also will have an impact on this algorithm. These two parameters are normalized between $[0:1]$. α is a parameter that can be changed, depending on the database that we are use. If $\alpha = 1$ we are in the method proposed by Scheidat et al. (2007). Through this approach it is expected that the algorithm will be able to fit better to the data, and will update with more equilibrium.

$$Si = \alpha \times d(actual, i) + (1 - \alpha) \times antiquity \tag{7.1}$$

7.5 Results and Discussion

The performed experiments aim to evaluate the need of template update, even as impact of those techniques. In Fig. 7.2, 7.3, 7.4 and 7.5 we present diverse results regarding several thresholds in

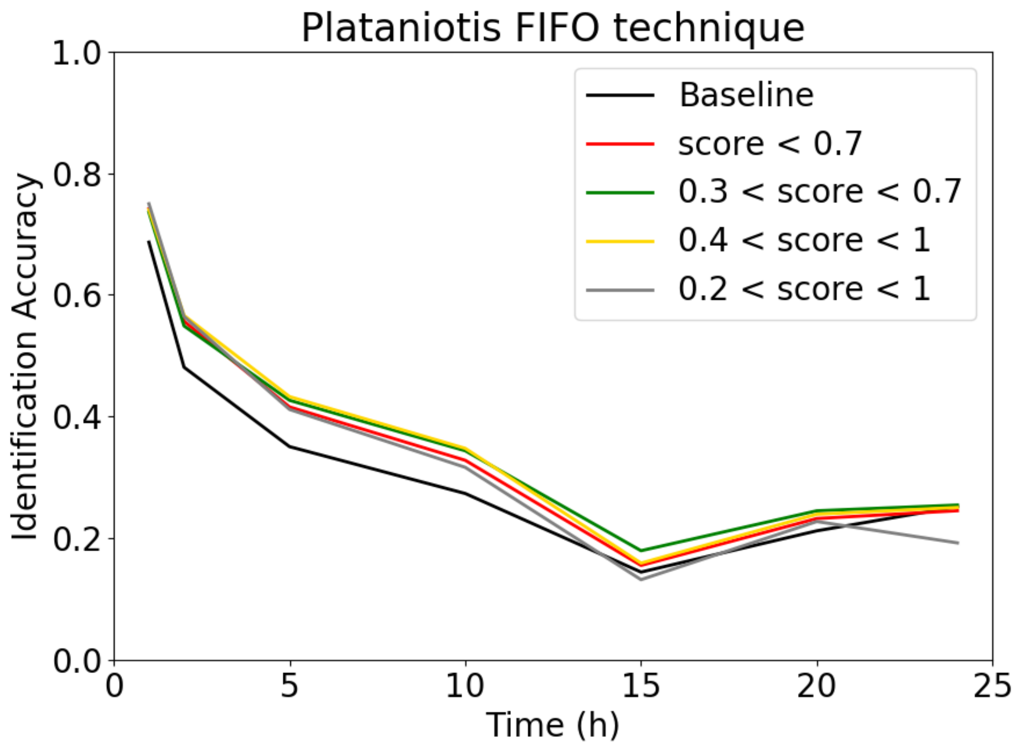


Figure 7.2: Comparative FIFO methods with different thresholds in Plataniotis method.

order to choose the most suitable for our purpose. According, we choose for the next steps in Plataniotis a threshold where the distance between samples varies $[0.3, 0.7]$. For the Auto-encoder we choose a threshold were the distance between samples varies $[0.1, 0.3]$. In the Belgacem method, the results improve in the first two hour (threshold for the difference between the highest score and the second highest stays in $[0.15, 0.3]$) but than decay significantly. In the Tawfik method the same is verify (the threshold for a highest improve in the first two hours is bigger than 0.2). In cNN method we try with through backpropagation update the weights but didn't result (see Fig. 7.6). The better results prove to be the ones with a range, near to the example shown before in section 7.2.1 (there was a need to reproduce a similar graph for each method).

With the results obtain, we suggest that classifiers as Random Forest, MLP and with the last layer in cNN (softmax) are not suitable for these kind of template/model update, besides the fact that improve in the beginning, the performance decay significantly over time. As such, in order to better understand this idea we decide to implement and replace these classifiers and the last layer of cNN with kNN and the results are shown in Fig. 7.8, Fig. 7.7 and Fig. 7.9. These results support our idea.

After, testing and decide the value for the threshold with FIFO technique, we proceed for Fixation technique and we obtain promising results, namely in $j \times 3 + 3$ approach as can be seen in Fig. 7.10, Fig. 7.11, Fig. 7.12 and in Fig. 7.13. This might be because the model learns better the ECG variability.

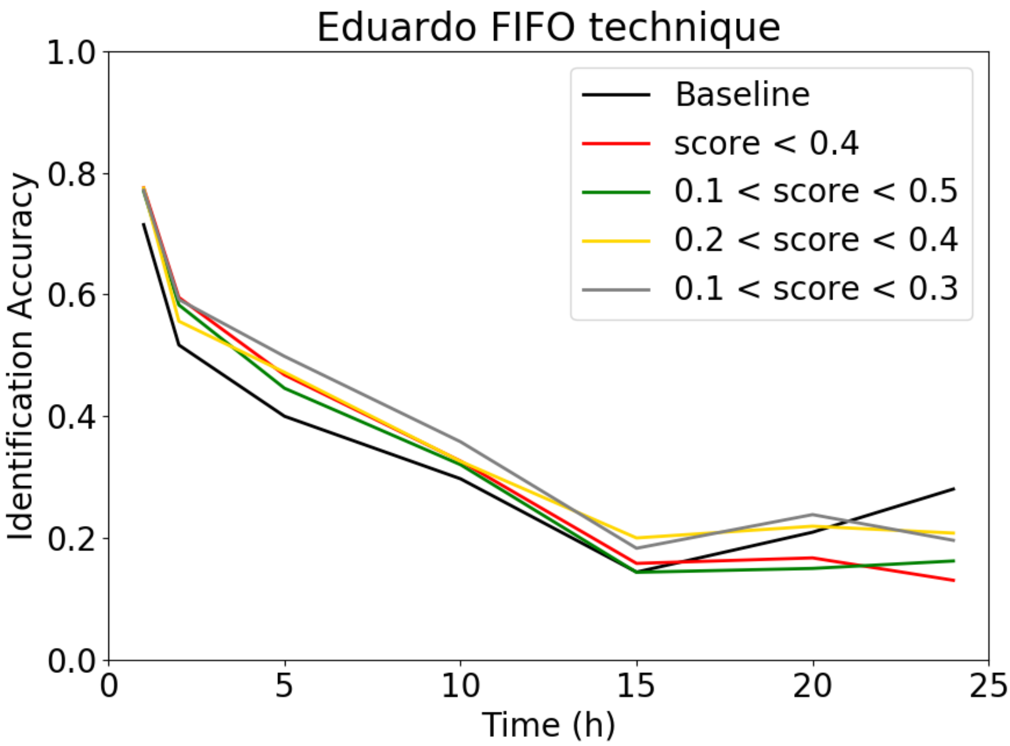


Figure 7.3: Comparative FIFO methods with different thresholds in Auto-encoder method.

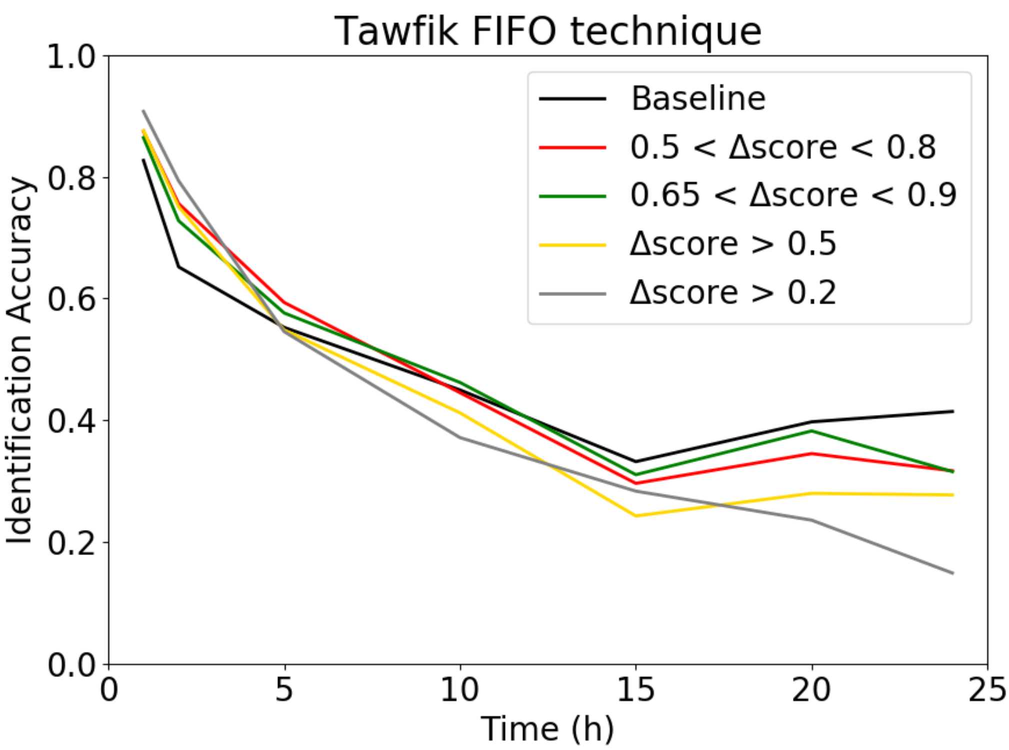


Figure 7.4: Comparative FIFO methods with different thresholds in Tawfik method.

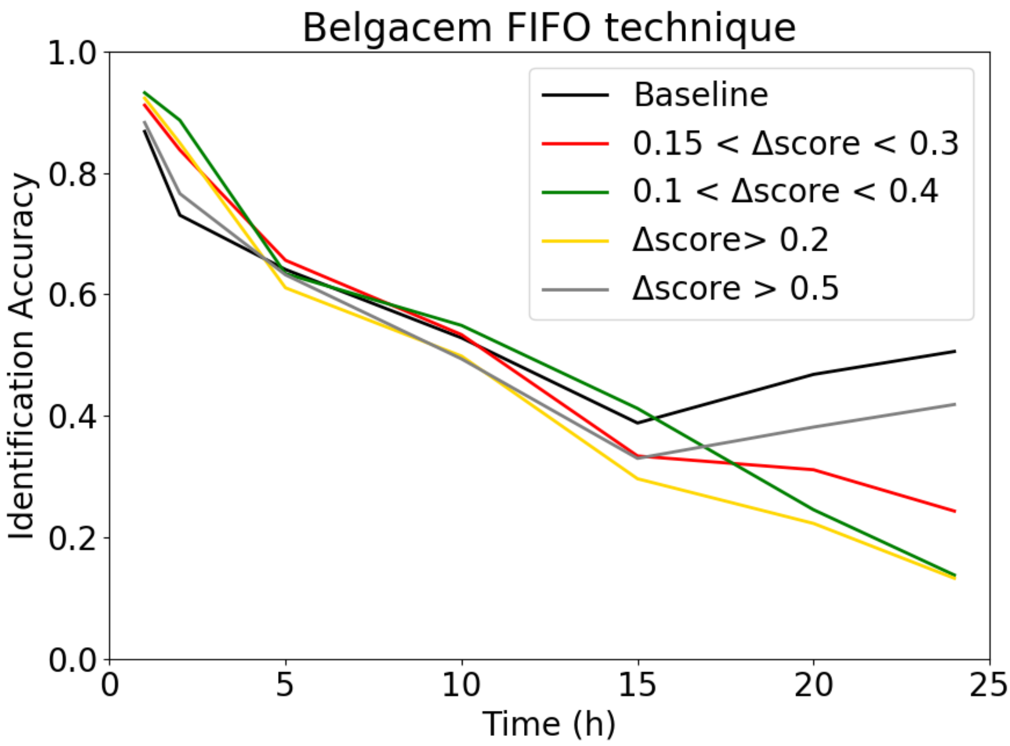


Figure 7.5: Comparative FIFO methods with different thresholds in Belgacem method.

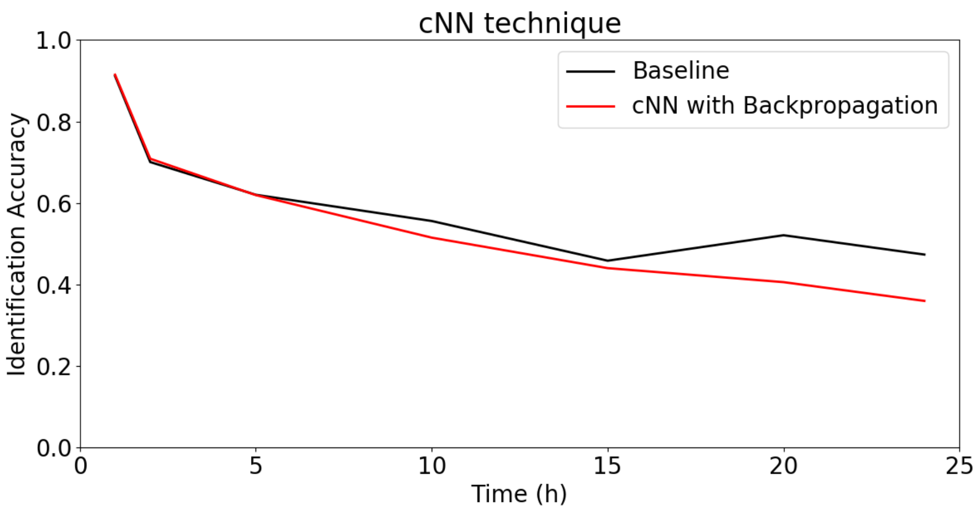


Figure 7.6: Comparative cNN method vs cNN with backpropagation.

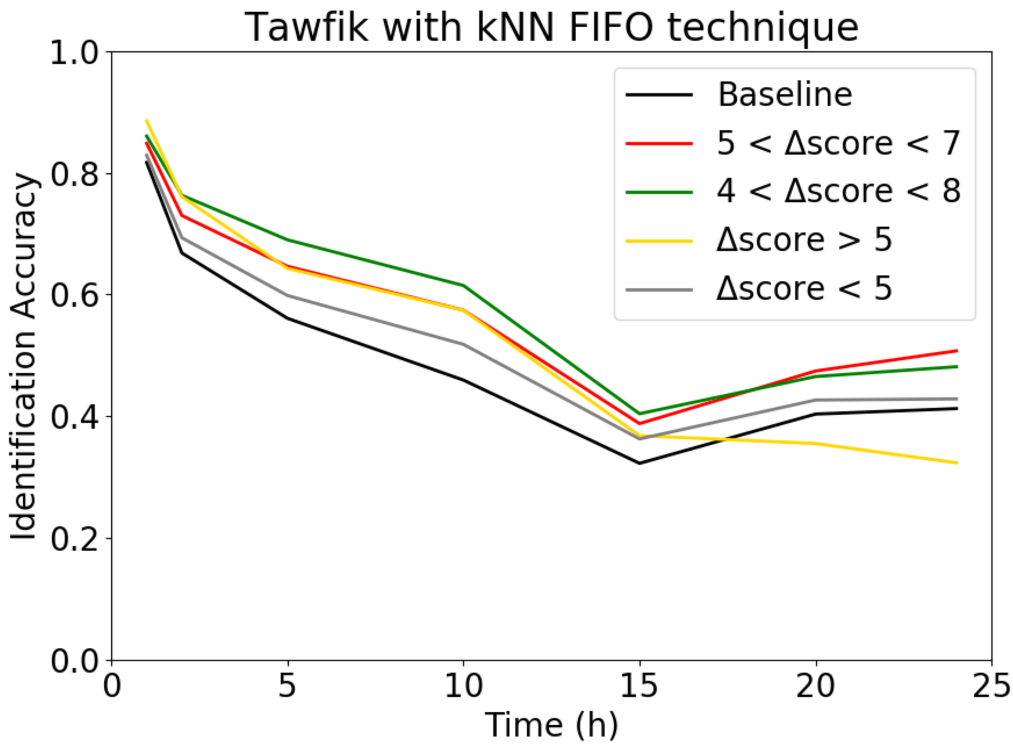


Figure 7.7: Comparative FIFO methods with different thresholds in Tawfik method with kNN.

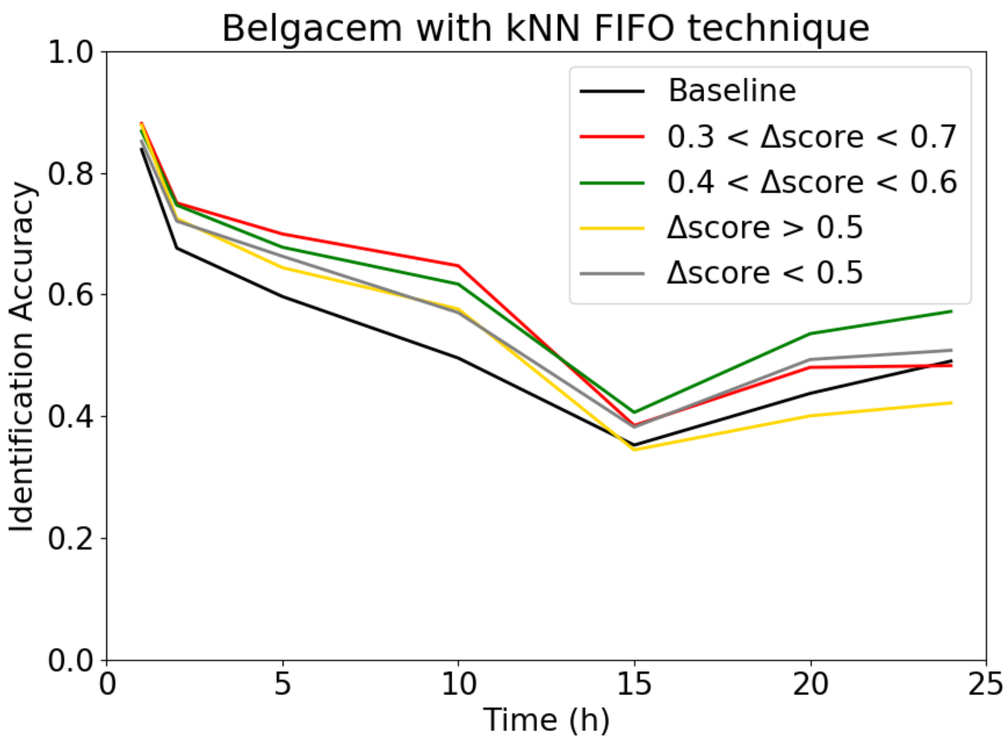


Figure 7.8: Comparative FIFO methods with different thresholds in Belgacem method with kNN.

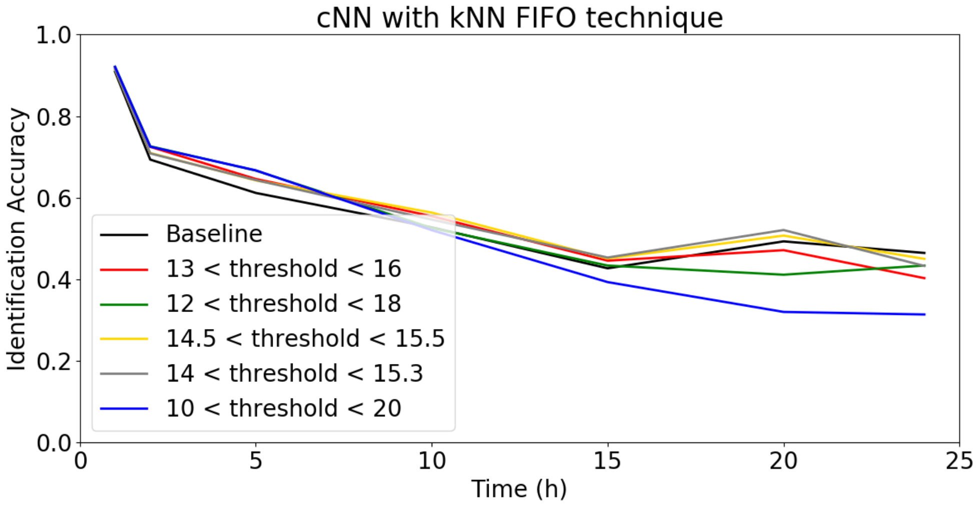


Figure 7.9: Comparative FIFO methods with different thresholds in cNN method with kNN.

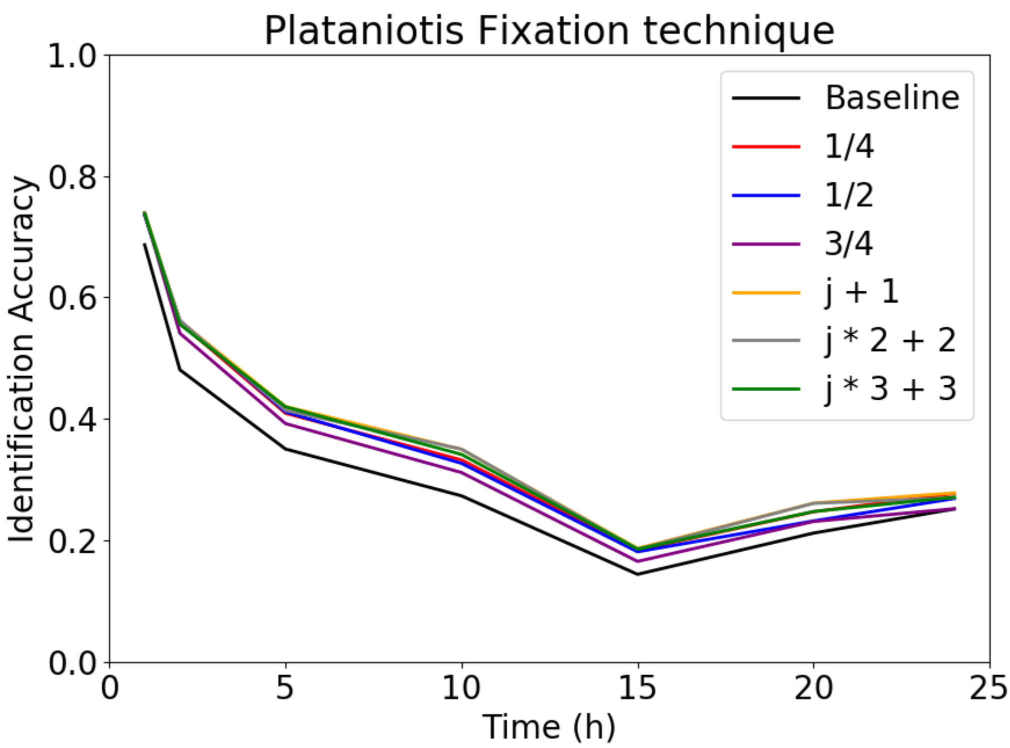


Figure 7.10: Comparative Fixation methods.

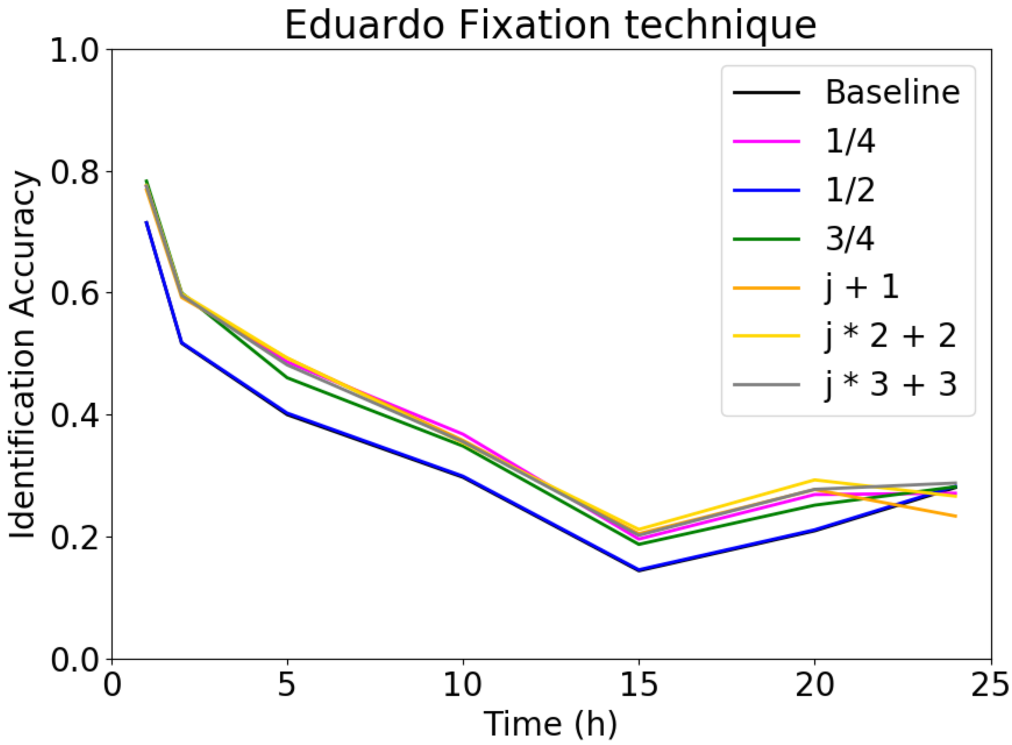


Figure 7.11: Comparative Fixation methods .

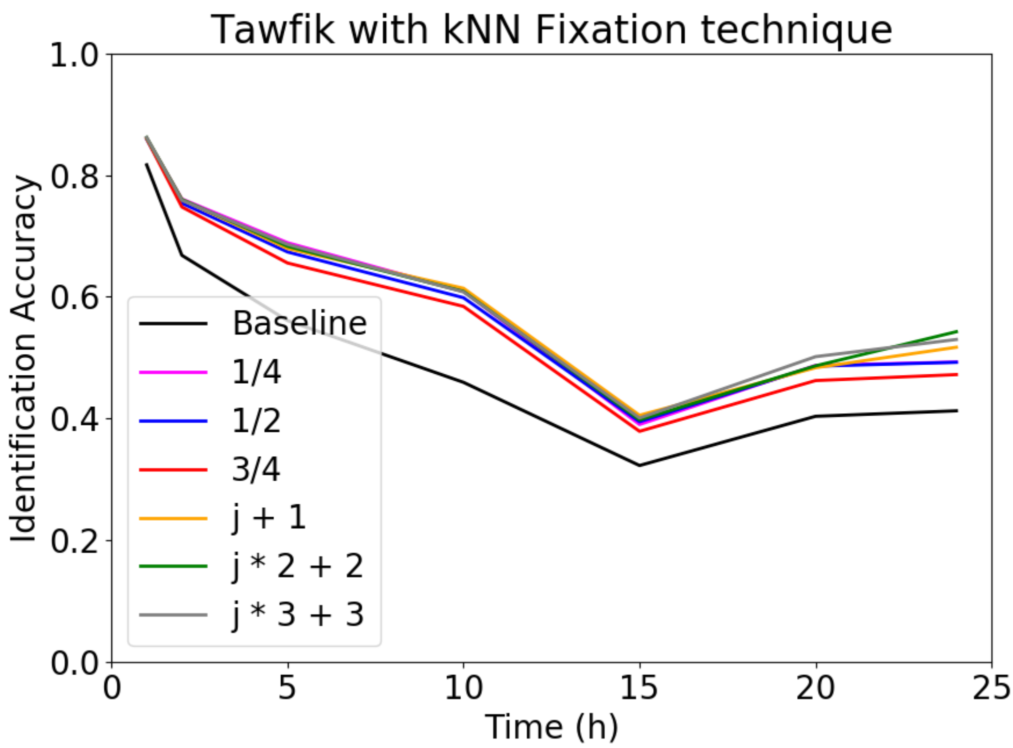


Figure 7.12: Comparative Fixation methods with different thresholds in Tawfik method with kNN.

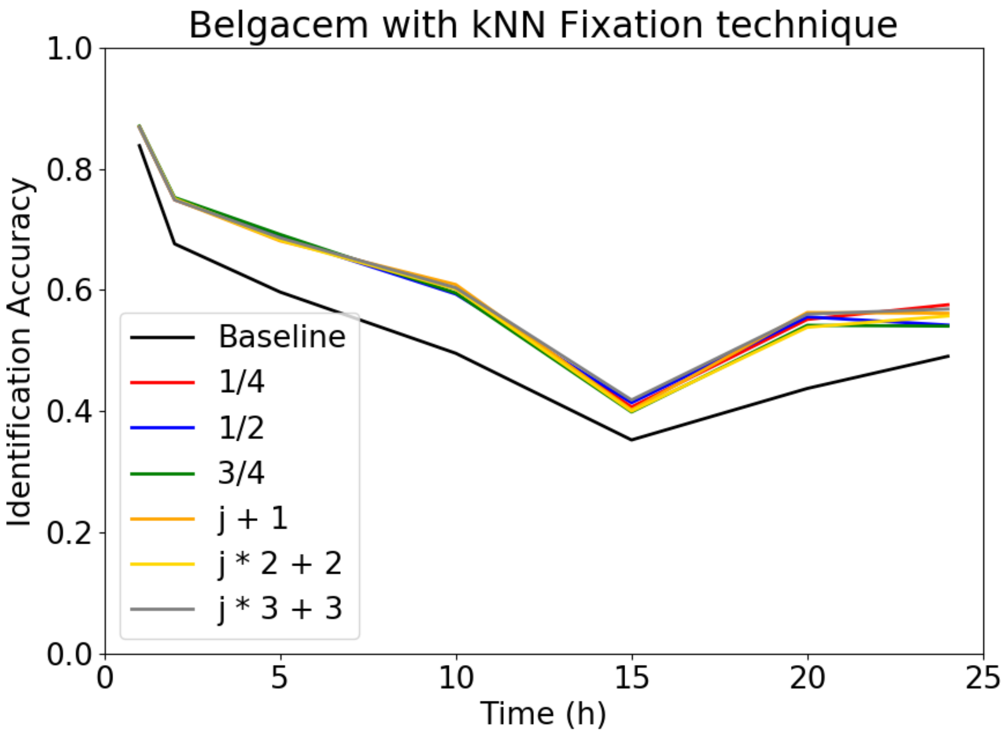


Figure 7.13: Comparative Fixation methods with different thresholds in Belgacem method with kNN.

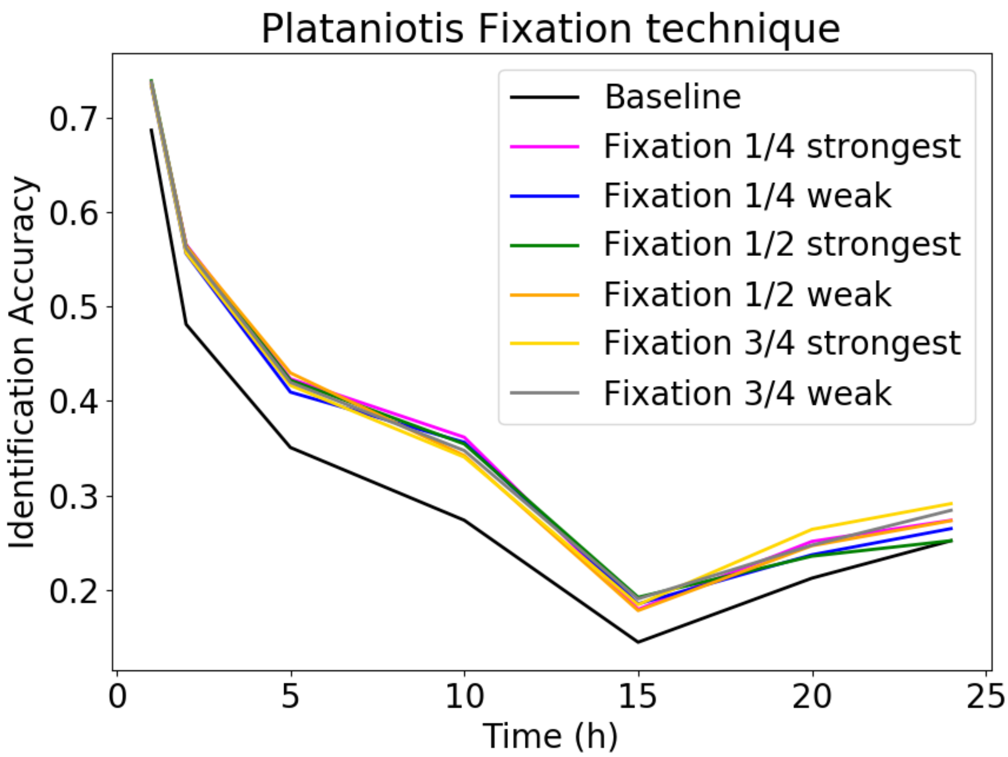


Figure 7.14: Comparative Fixation methods with the fixation of the strongest and the weakest samples.

Instead of the normal Fixation technique, this method allows the samples to compete for reselection, witch is very similar to the clock method presented in Chapter 5. The samples are chosen randomly with the correlation matrix the performance improve significantly comparative to the baseline but didn't show changes in the results comparative to the random fixation as can be seen in Fig. 7.14.

The method where we obtain better improvements with our techniques is in the tawfik method, that we where able of improve the mean of results in 10%.

Chapter 8

Conclusion

8.1 Final Remarks and Future Work

The work explained before, focus on testing the impact of *template update* in a realistic environment for ECG biometrics in a continuous *long-term* identification measurement in order to mitigate performance decay over time. To accomplish that, it was exploit exhaustively 5 state-of-the-art methods with two main template update techniques and a propose adaptive technique. Also study how the ECG variability effects the performance of state-of-the-art biometrics algorithms. The results reveal ECG signals are unreliable for *long-term* biometric applications, despite the promising results often presented in the literature.

Template update techniques proved successful in enhancing the long-term performance of state-of-the-art methods, especially when using template fixation techniques. However, further efforts are needed for the study and development of more advanced techniques, with special focus on supervised techniques, so that ECG-based biometric systems can offer reliable performances over long periods of operation. The rate of accepted templates (see Appendix) shown to be not that much significant and a different study can be performed in the future, mainly if, it was mandatory an *x%* of templates to be replace at each test instance (i.e other rule for the update, instead of the threshold). Other considerations, should be when it is done the update (since we have done batch-batch, other interesting idea it was study the impact of doing in two in two batch and so on). One different study that can be made, correspond to experiment other remaining classifiers to better understand which are suitable or not for the diverse *template update* methodologies.

To conclude, other approach could be of the most interest, Extremely Fast Decision Tree (EFDT) proposed by Manapragada et al. (2018) (appears as an alternative to the traditional classifiers). The authors defend that this classifier has statistically advantages. The concept of Drift adaptation is exploit, and generally requires forgetting mechanisms for update the model. With this in mind and since that is similar with the techniques exploit during this dissertation, there exist space for an adaptation of this "dynamic model" to ECG *long-term* biometric application.

Appendix A

Appendix

In this chapter, is presented a summary of all the results shown in the charts of the previous Chapters, including the template acceptance rate at each hour.

Table A.1: Plataniotis results with different starting times.

Start Time:	**x_test_1h:**	**x_test_2h:**	**x_test_5h:**	**x_test_10h:**	**x_test_15h:**	**x_test_20h:**	**x_test_24h:**
03:00	84,79	60,59	23,23	11,87	12,64	27,6	62,36
08:00	69,37	45,04	25,72	20,52	20,58	15,87	38,34
12:00	70,86	49,41	30,56	22,4	10,38	22,71	39,27
16:00	67,6	40,81	32,19	15,59	19,78	29,36	48,61

Table A.2: Comparison between all the techniques with Plataniotis method. FIFO_1 represents threshold ≥ 0.7; FIFO_2 represents $0.3 \leq$ threshold ≤ 0.7; FIFO_3 represents $0.4 \leq$ threshold ≤ 1; FIFO_4 represents $0.2 \leq$ threshold ≤ 1.

Technique	Test_1h	%Update	Test_2h	%Update	Test_5h	%Update	Test_10h	%Update	Test_15h	%Update	Test_20h	%Update	Test_24h	%Update
Normal:	68,65	-	48,10	-	35,06	-	27,36	-	14,45	-	21,23	-	25,20	-
FIFO_1	74,13	61,72	55,66	45,67	41,59	38,50	32,82	32,99	15,55	20,56	23,25	21,19	24,54	20,75
FIFO_2	73,83	38,53	55,64	30,57	41,78	27,02	35,27	21,12	18,11	10,16	24,52	14,51	27,75	15,27
FIFO_2 more similar	73,69	39,65	55,39	28,92	42,26	25,88	35,42	21,82	18,67	10,43	23,83	15,58	25,20	17,25
FIFO_2 less similar	73,56	39,65	54,89	28,92	42,26	25,88	34,39	21,82	17,96	10,43	24,52	15,58	25,46	17,25
FIFO_3	74,01	40,44	56,61	41,52	43,28	40,16	34,78	36,88	15,93	23,14	23,96	31,90	25,07	29,71
FIFO_4	74,96	69,57	56,44	64,73	41,18	62,66	31,66	55,74	13,21	38,07	22,76	46,51	19,25	41,07
Fixation n/4	73,67	38,35	55,85	31,06	40,94	25,94	33,26	20,81	18,43	10,2	24,74	14,60	27,75	16,92
Fixation n/4 more similar	73,58	38,72	56,54	31,17	42,34	27,50	36,15	21,21	17,93	11,35	25,12	15,69	27,36	20,51
Fixation n/4 less similar	73,39	39,24	55,60	31,13	40,93	28,16	35,62	23,48	18,50	11,48	23,70	15,40	26,48	18,53
Fixation n/2	73,49	38,88	56,21	30,90	41,16	25,24	32,71	20,48	18,16	9,96	23,23	12,91	26,92	17,39
Fixation n/2 more similar	73,88	38,90	56,42	31,17	42,19	26,88	35,44	22,09	19,21	12,09	23,56	15,80	25,20	18,85
Fixation n/2 less similar	73,67	38,64	56,38	30,53	42,96	26,65	34,21	22,17	17,80	11,15	24,67	15,14	27,31	17,12
Fixation 3n/4	73,9	39,13	54,11	30,04	39,25	23,62	31,19	18,48	16,58	10,11	23,14	12,36	25,24	14,55
Fixation 3n/4 more similar	73,46	38,99	55,66	31,42	41,64	26,92	34,05	21,09	18,43	11,12	26,4	16,09	29,12	17,47
Fixation 3n/4 less similar	73,74	39,2	56,15	30,89	41,96	27,5	34,74	21,93	19,05	11,57	24,72	16,14	28,41	19,7
Fixation j*1+1	73,95	38,4	56,12	30,96	42,07	26,47	35,01	21,5	18,72	11,01	26,16	15,53	27,84	18,2
Fixation j*2+2	73,92	38,42	56,13	31,12	41,43	26,99	35,04	22,17	18,58	11,48	26,12	15,69	27,05	17,86
Fixation j*3+3	73,72	38,48	55,57	31,12	42,00	27,11	34,17	21,94	18,59	11,03	24,78	15,43	27,14	18,47
Clock min distance	68,01	37,59	37,52	19,17	34,26	10,57	26,6	6,45	14,02	4,4	20,6	6,92	24,23	6,43
Clock max distance	67,87	38,72	47,13	23,26	32,64	14,08	25,12	10,04	13,96	7,69	18,6	10,05	22,73	9,83
Clock (alpha=0.5 and min)	67,29	35,5	39,91	11,17	28,15	5,44	19,34	2,62	11,15	2,02	15,92	2,63	16,87	1,9
Clock (alpha=0.5 and max)	67,85	38,3	47,23	22,82	33,83	14,4	26,97	10,21	14,22	8,23	20,01	10,87	22,56	11,17
Clock (alpha=0.9 and max)	68,01	38,16	46,74	20,32	32,75	11,12	24,97	7,22	13,53	5,32	18,52	7,49	22,33	7,29
Clock (alpha=0.9 and min)	67,43	36,01	41,97	13,03	29,71	6,15	20,78	3,37	12,13	2,38	15,78	3,49	17,14	2,99
Clock (alpha=0,1 and max)	68,12	38,65	47,71	23,9	34,96	15,49	27,31	11,12	14,68	8,9	20,56	11,2	24,05	11,85
Clock (alpha=0.1 and min)	66,72	34,8	38,1	10,71	26,49	4,87	17,61	2,6	10,63	1,91	14,22	2,63	15,24	1,86
Clock (alpha=0.25 and max)	67,93	38,49	47,41	23,23	34,49	14,99	26,93	10,66	14,27	8,34	20,51	11,6	23,26	11,52
Clock (alpha=0.25 and min)	66,68	34,77	38,35	11,01	27,08	4,88	17,86	2,62	10,86	1,95	14,96	2,67	16,12	1,61
Clock (alpha=0.75 and max)	67,96	38,56	46,79	21,71	33,71	12,92	26,08	8,65	14,14	6,79	19,01	8,07	22,29	8,20
Clock (alpha=0.75 and min)	67,00	35,71	40,83	11,33	28,59	05,65	19,88	2,82	11,14	2,31	15,63	3,23	18,72	2,31

Table A.3: Comparison between all the techniques with Eduardo method. FIFO_1 represents threshold ≤ 0.4; FIFO_2 represents $0.1 \leq$ threshold ≤ 0.5; FIFO_3 represents $0.2 \leq$ threshold ≤ 0.4; FIFO_4 represents $0.1 \leq$ threshold ≤ 0.3.

Method	Technique	Test_1h	%Update	Test_2h	%Update	Test_5h	%Update	Test_10h	%Update	Test_15h	%Update	Test_20h	%Update	Test_24h	%Update
Eduardo	**Normal:**	71,47	-	51,72	-	40,02	-	29,75	-	14,43	-	20,98	-	28,06	-
	FIFO_1	77,52	79,61	59,52	76,97	46,81	76,79	32,66	73,14	15,86	66,25	16,76	69,21	13,08	55,89
	FIFO_2	76,93	78,63	58,3	82,98	44,62	78,72	32,07	73,74	14,4	60,49	15,03	69,69	16,26	70,08
	FIFO_3	77,46	33,05	55,62	35,76	47,24	34,17	32,62	32,19	20,04	23,71	21,96	31,29	20,84	35,01
	FIFO_4	77,13	60,99	59,18	61,52	49,86	57,15	35,83	50,8	18,34	27,44	23,87	40,98	19,65	40,8
	Fixation n/4	77,43	61,42	59,75	61,15	48,66	55,81	36,81	50,36	19,62	29,19	26,92	41,51	27,14	41,1
	Fixation n/2	77,73	61,12	61,21	60,12	48,77	55,79	34,55	49,09	20,07	27,89	26,25	38,02	,28,02	41,91
	Fixation 3n/4	78,28	61,52	59,96	58,85	46,06	53,81	34,87	46,17	18,76	27,42	25,18	35,87	28,19	41,44
	Fixation j*1+1	76,86	61,17	59,22	61,58	49,27	57,65	35,79	51,93	20,5	29,73	27,74	41,31	23,39	43,62
	Fixation j*2+2	77,59	61,21	59,82	61,68	49,3	56,76	35,38	51,68	21,21	29,86	29,32	43,03	26,65	44,29
	Fixation j*3+3	77,46	61,54	59,57	61,65	48,16	57,58	35,58	50,62	20,25	32,5	27,78	44,96	28,81	43,83

Table A.4: Comparison between all the techniques with Belgacem method. In the traditional technique FIFO_1 represents $0.15 \leq$ threshold ≤ 0.3; FIFO_2 represents $0.1 \leq$ threshold ≤ 0.4; FIFO_3 represents threshold ≥ 0.2; FIFO_4 represents threshold ≥ 0.5. In the kNN Belgacem FIFO_1 represents $0.3 \leq$ threshold ≤ 0.7; FIFO_2 represents $0.4 \leq$ threshold ≤ 0.6; FIFO_3 represents threshold ≥ 0.5; FIFO_4 represents threshold ≤ 0.5.

Method	**Technique**	**Test_1h**	**%Update**	**Test_2h**	**%Update**	**Test_5h**	**%Update**	**Test_10h**	**%Update**	**Test_15h**	**%Update**	**Test_20h**	**%Update**	**Test_24h**	**%Update**
Belgacem:	**Normal:**	86,81	-	73,03	-	64,1	-	52,85	-	38,85	-	46,85	-	50,62	-
	FIFO_1	91,13	9,8	83,76	12,18	65,61	16,88	53,39	18,2	33,42	15,03	31,16	18,7	24,36	19,49
	FIFO_2	93,17	19,27	88,65	23,81	63,4	33,74	54,9	33,96	41,24	29,86	24,58	37,66	13,83	36,77
	FIFO_3	92,29	87,82	,84,91	78,81	61,09	55,49	49,86	47,7	29,68	35,71	22,32	25,46	13,3	12,82
	FIFO_4	88,28	65,04	76,58	50,14	63,23	22,83	49,39	14,94	33,03	8,16	38,17	4,07	41,89	3,41
Belgacem_knn	**Normal:**	83,81	-	67,62	-	59,68	-	49,57	-	35,28	-	43,77	-	49,07	-
	FIFO_1	88,09	44,72	75,04	43,9	69,95	41,5	64,72	35,78	38,5	20,27	48,05	33,43	48,33	32,5
	FIFO_2	86,84	20,32	74,68	20,48	67,77	18,34	61,71	16,08	40,65	8,72	53,57	14,78	57,22	16,26
	FIFO_3	87,82	22,46	72,46	24,79	64,42	23,9	57,63	23,26	34,49	20,02	40,08	25,06	42,2	33,02
	FIFO_4	85,16	68,76	72,06	57,94	66,29	46,83	57,02	37,93	38,23	18,05	49,34	25,81	50,84	27,8
	Fixation n/4	86,92	20,16	74,89	20,6	68,63	18,32	60,23	16,01	40,72	8,5	55,15	15,38	57,58	16,48
	Fixation n/2	86,83	20,43	75,11	21,1	68,72	18,31	59,34	15,95	41,42	9,17	55,57	15,19	54,23	14,95
	Fixation 3n/4	87,04	20,44	75,27	21,74	69,18	18,11	59,52	15,78	39,89	8,9	54,19	13,67	54,05	12,98
	Fixation j*1+1	86,9	20,28	74,93	20,69	68,11	18,4	60,91	16,17	40,25	8,86	56,3	15,23	56,17	16,15
	Fixation j*2+2	,86,88	20,46	75,07	20,78	68,25	18,43	60,16	16,49	40,04	8,95	53,86	15,39	55,73	15,67
	Fixation j*3+3	86,91	20,09	74,88	20,67	68,61	18,36	60,37	16,4	41,87	9,08	56,04	15,61	56,87	16,83

Table A.5: Comparison between all the techniques with Tawfik method. In the traditional technique FIFO_1 represents $0.5 \leq$ threshold ≤ 0.8; FIFO_2 represents $0.65 \leq$ threshold ≤ 0.9; FIFO_3 represents threshold ≥ 0.5; FIFO_4 represents threshold ≤ 0.5. In the kNN Tawfik FIFO_1 represents $5 \leq$ threshold ≤ 7; FIFO_2 represents $4 \leq$ threshold ≤ 8; FIFO_3 represents threshold ≥ 5; FIFO_4 represents threshold ≤ 5.

Method	**Technique**	**Test_1h**	**%Update**	**Test_2h**	**%Update**	**Test_5h**	**%Update**	**Test_10h**	**%Update**	**Test_15h**	**%Update**	**Test_20h**	**%Update**	**Test_24h**	**%Update**
Tawfik:	**Normal:**	82,68	-	65,2	-	55,26	-	44,99	-	33,23	-	39,76	-	41,45	-
	FIFO_1	87,46	15,87	75,59	19,59	59,36	16,79	44,15	11,6	29,66	8,49	34,54	7,67	31,72	6,41
	FIFO_2	86,42	23,1	72,77	23,63	57,63	14,72	46,22	10,41	31,08	6,49	38,27	5,96	31,59	3,81
	FIFO_3	87,57	81,47	75,07	63,88	54,85	31,96	41,23	21,52	24,31	12,99	28,01	6,9	27,75	5,55
	FIFO_4	90,73	92,78	79,31	83,34	54,56	53,23	37,17	39,64	28,38	33,08	23,61	18,93	14,93	13,86
Tawfik_knn:	**Normal:**	81,68	-	66,83	-	56,1	-	45,95	-	32,29	-	40,37	-	41,28	-
	FIFO_1	84,84	18,94	72,98	16,76	64,63	13,46	57,43	10,94	38,81	5,1	47,45	8,56	50,75	10,13
	FIFO_2	86,01	38,28	76,29	36,01	68,98	32,28	61,48	26,31	40,43	12,34	46,54	20,8	48,15	20,52
	FIFO_3	88,53	45,78	76,06	48,37	64,28	45,29	57,42	42,66	36,83	34,00	35,54	43,87	32,38	50,75
	FIFO_4	82,87	39,54	69,33	22,46	59,84	16,83	51,82	9,48	36,92	2,16	42,68	5,00	42,86	7,79
	Fixation n/4	86,08	38,55	76,08	36,45	68,89	31,68	60,91	25,81	39,03	11,3	48,74	20,61	49,3	22,55
	Fixation n/2	85,92	38,51	75,39	36,05	67,38	31,05	59,86	24,28	39,48	10,43	48,61	19,25	49,25	20,84
	Fixation 3n/4	85,98	39,02	74,77	35,78	65,56	29,39	58,43	23,39	37,91	10,2	46,25	18,04	47,22	18,54
	Fixation j*1+1	86,08	38,32	75,98	35,89	67,93	32,14	61,41	26,19	40,54	12,34	48,35	21,56	51,72	22,21
	Fixation j*2+2	86,17	38,44	75,99	35,8	68,31	32,26	60,89	25,81	39,78	11,87	48,66	21,78	54,27	22,54
	Fixation j*3+3	86,19	38,55	75,87	35,85	68,63	32,07	60,8	26,221	40,09	0,12,07	50,17	21,07	53,00	22,50

Table A.6: Comparison between all the techniques with cNN. FIFO_1 represents 13 $\leq$ threshold $\leq$ 16; FIFO_2 represents 12 $\leq$ threshold $\leq$ 18; FIFO_3 represents 14.5 $\leq$ threshold $\leq$ 15.5; FIFO_4 represents 14 $\leq$ threshold $\leq$ 15.3; FIFO_5 represents 10 $\leq$ threshold $\leq$ 12.

Method	**Technique**	**Test_1h**	**%Update**	**Test_2h**	**%Update**	**Test_5h**	**%Update**	**Test_10h**	**%Update**	**Test_15h**	**%Update**	**Test_20h**	**%Update**	**Test_24h**	**%Update**
Cnn	**Normal**	91,15	-	70,03	-	62,00	-	55,56	-	45,82	-	52,07	-	47,32	-
	Back_Propagation	91,48	-	70,86	-	61,92	-	51,48	-	43,99	-	40,55	-	35,95	-
Cnn_knn	**Normal**	90,89	-	69,33	-	61,15	-	52,72	-	42,66	-	49,27	-	46,44	-
	FIFO_1	91,88	7,04	72,46	6,31	64,58	5,27	55,46	4,63	44,53	3,23	47,11	3,68	40,25	5,81
	FIFO_2	92,1	12,1	72,6	12,1	66,64	10,44	52,56	9,96	43,34	8,53	41,09	9,04	43,35	9,52
	FIFO_3	91,46	2,22	70,94	1,86	64,23	1,52	56,33	1,32	45,07	0,77	50,68	1,26	44,98	1,54
	FIFO_4	91,48	3,06	70,85	2,31	64,31	2,09	54,62	1,73	45,31	1,01	52,05	1,5	43,25	2,32
	FIFO_5	92,07	16,67	72,5	16,41	66,69	15,91	52,05	14,76	39,28	14,00	31,97	14,41	31,34	15,26

Don't You Forget About Me: A Study on Long-Term Performance in ECG Biometrics

Gabriel Lopes[1], João Ribeiro Pinto[1,2][0000−0003−4956−5902], and Jaime S. Cardoso[1,2][0000−0002−3760−2473]

[1] Faculdade de Engenharia, Universidade do Porto, Porto, Portugal
[2] Centre for Telecommunications and Multimedia, INESC TEC, Porto, Portugal

Corresponding author: João Ribeiro Pinto (jtpinto@fe.up.pt)

Abstract. The performance of biometric systems is known to decay over time, eventually rendering them ineffective. Focused on ECG-based biometrics, this work aims to study the permanence of these signals for biometric identification in state-of-the-art methods, and measure the effect of template update on their long-term performance. Ensuring realistic testing settings, four literature methods based on autocorrelation, autoencoders, and discrete wavelet and cosine transforms, were evaluated with and without template update, using Holter signals from THEW's E-HOL 24h database. The results reveal ECG signals are unreliable for long-term biometric applications, and template update techniques offer considerable improvements over the state-of-the-art results. Nevertheless, further efforts are required to ensure long-term effectiveness in real applications.

Keywords: Biometrics · Electrocardiogram · Identification · Template Update

1 Introduction

The ability to identify or recognize another human being is of utmost importance. For access control and other purposes, security systems are based on unique credentials. Several methods are used such as PIN codes, passwords, ID cards, or keys [1]. These credentials are susceptible to be copied or counterfeit [2], and it is very easy for the person to forget, share, or lose them. The alternative is for systems to use something that is characteristic and belongs only to one person: a biometric trait [3]. Biometric traits currently include fingerprint, voice, face, and electrocardiogram (ECG), and only require the individual to be present when access is requested [4]. Relative to the traditional techniques, biometric systems are considered more secure, as traits are more difficult to counterfeited or steal than extrinsic credentials [5].

Current off-the-person ECG acquisition techniques, aiming towards increased simplicity, usability, and comfort, make ECG-based biometric systems effortless for the user [6, 7]. However, when considering long-term usage, the performance decays over time [8]. This applies to human-machine interfaces that require frequent identity control, especially those that are information-sensitive.

Variability on the input biometric data, biometric trait's aging, and variations caused by the subject's interaction with the sensor contribute to large intrasubject variability [5]. This makes stored individual templates to quickly lose representativity, resulting in poor recognition performance and placing serious challenges on long-term recognition. In biometrics in general, long-term identification requires frequent update of the templates to maintain acceptable performance over time [9], and avoid security flaws in information-sensitive systems.

Previously, Labati *et al.* [10] have studied the performance decay over time on their proposed algorithm for ECG-based authentication. In this paper, we extend their work, by studying the effect of ECG permanence and variability in biometric identification and evaluating the impact of template update on long-term performance in four state-of-the-art algorithms for identification.

2 State-of-the-art review

In the literature, it is difficult to find a strong and widely accepted rule for template update. Most methods are based on heuristics and empirically determined thresholds, which are highly dependent on the data and application settings. For example, Komeili *et al.* [8], for authentication, have set the acceptance threshold equal to the point of zero false acceptance rate, thus ensuring updates with only genuine samples.

Nevertheless, it is possible to identify some common mechanisms that may vary depending on different factors: these include the choice of the update criterion (based on thresholds or graphs), the update periodicity (online or offline), the selection mechanism, and the template update working mode system (supervised or semi-supervised). The taxonomy of template update (see Fig. 1) divides the existing techniques into two categories: supervised and semi-supervised.

Supervised methods are offline methods in which label attribution is given by a supervisor. These contain the *Clustering* subcategory, which includes the *MDIST* that aims to search for the templates that minimize the intra distance among all the samples in the database (*i.e.*, the most similar) and *DEND* that aims to search for the templates that exhibit large intra-class variations resort to the dendrogram (*i.e.*, the most different) [12]. The second subcategory comprises Editing-based methods, which are independent of the number of templates and give focus on the whole collected training set T. A subset $E \in T$ is generated, maintaining the classification performance offered by T. The best subsets were obtained by reviewing the structure of the data (needs to be done to each subject) [9, 13]. All the algorithms (based on k-Nearest Neighbors) must be representative of T and can be roughly described as *incremental* when the E starts empty and grows, or *decremental* when E starts equal to T and in each iteration some instances are deleted, until some criterion is reached [9].

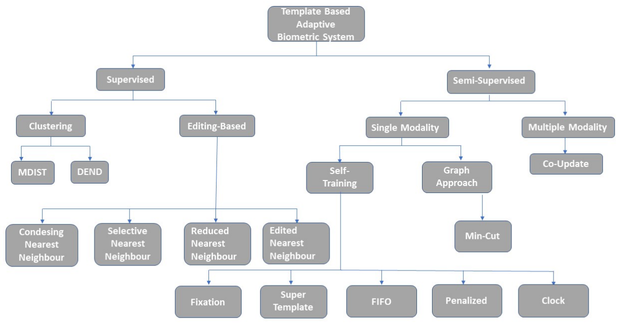


Fig. 1: Dendrogram representing the taxonomy of template update techniques (based on [11]).

Semi-supervised methods merge labeled (in biometrics correspond to the initial training samples) and unlabeled (correspond to the samples available during system operation) data to improve the system's performance. This category comprises the *Single Modality* (for unimodal biometric systems) and *Multiple Modality* (for systems using more than one biometric trait) subcategories. The *Single Modality* subcategory includes the *Self-Training* approaches such as *FIFO* (first-in-first-out), *Fixation*, *Super Template* (X composed by N templates x) where new genuine date is always fused to a common single template [14] updated online during the execution of continuous verification, *Penalized template update* method based on the mean of the past ECG's and the actual ECG [15] and *clock method* where the current template is tested against all the others stored in the database [16].

Generally, a new, unknown trait measurement is used for template update if its score (returned by the biometric recognition system) is above a set threshold. Hence, the future performance of the system relies heavily on the chosen threshold value [11].

The update threshold is commonly estimated using enrollment templates or training data. When training data are scarce or when using short enrollments, this leads to some problems: as important intrasubject variability is missed because use only the patterns similar to the templates stored; online methods are dependent on the order of the sequence of input data; vulnerability to large intra-class variations; since the algorithm normally looks for the minimal cost (high score), it can be stuck in local maximum and always use highly confident data for updating.

Table 1: Graph-based template update methods and their respective loss and regularizer functions (based on [17]).

Method	Source	Loss	Regularizer
Min Cut	[18]	$\sum_{i \in L}(y_i - y_{i\|L})^2$	$\frac{1}{2}\sum_{ij} w_{ij}(y_i - y_j)^2$
Gaussian Random Fields and harmonic Function	[19]	$\sum_{i \in L}(f_i - y_i)^2$	$f^T \Delta f$
Local and Global Consistency	[20]	$\sum_{i=1}^{n}(f_i - y_i)^2$	$D^{-\frac{1}{2}} \Delta D^{\frac{1}{2}}$
Tikhonov Regularization	[21]	$\frac{1}{K}\sum_{i}(f_i - y_i)^2$	$\gamma f^T S f$
Manifold Regularization	[22]	$\frac{1}{l}\sum_{i=1}^{l} V(x_i, Y_i, f)$	$\gamma_A \\|f\\|_k^2 + \gamma_I \\|f\\|_I^2$
Graph Kernel from the Spectrum of Laplacian	[23]	$\min \frac{1}{2} w^T W$	$\exp(-\frac{\sigma}{2}\lambda)$
Spectral Graph Transducer	[19]	$\min c(f-\gamma)^T C(f-\gamma)$	$f^T L f$
Local Learning Regularization	[24]	$\min \frac{1}{k}\sum_{i=1}^{k}(y_i - f_k(x_i))^2$	$\frac{\gamma}{k}\\|f_k\\|^2$

Semi-supervised methods also include *Graph* approaches. These commonly define a graph where the nodes are either labeled (the identity is known) or unlabeled (unknown identity) data, and the edges (which can have different weights) are the similarity between those samples [11, 17]. To be considered a graph-based semi-supervised method, it must estimate a function f, approximate the known Y on the labeled nodes, and include two terms to turn the graph smooth: a loss function and a regularizer. These two terms are what define each approach (as can be seen in Table 1) [17], among which the most common in biometrics is min-cut graphs [11].

Considering the topic of template update is still to be adequately addressed on ECG biometrics, this work studies the effect of ECG permanence and variability in long-term identification performance. Furthermore, it aimed to evaluate the effect of template update techniques, on the performance of several state-of-the-art methods.

3 Methods

3.1 Implemented identification methods

To fully and objectively evaluate the effects of ECG variability on the performance of biometric algorithms, a study was conducted on four literature methods, described below.

Plataniotis *et al.* [25] proposed an ECG biometric recognition method using a non-fiducial approach. Signals are preprocessed using a bandpass filter ($0.5 - 40$ Hz), followed by feature extraction with autocorrelation (AC) and dimensionality reduction using discrete cosine transform (DCT). The fifteen most relevant features were selected, and Euclidean distance was used for classification.

Tawfik *et al.* [26] used a bandpass filter ($1-40$ Hz) in the preprocessing phase. QRS complexes (most stable part of ECG) were cut from the signal using a 0.35 second window. The average ensemble QRS was computed and features were extracted using DCT technique (thirty most relevant features were selected). A multilayer perceptron (MLP) is used for classification.

Belgacem *et al.* [27] also preprocessed signals with a bandpass filter ($1-40$ Hz). The QRS complexes were located and cut from the signal, and the average QRS was computed. The feature extraction resorted to Discrete Wavelet Transform (DWT). From all DWT decomposition levels, only the most relevant were selected, and a Random Forest is used for classification. The authors used this technique for authentication.

Eduardo *et al.* [28] used a Finite Impulse Response ($5-20$ Hz) filter for preprocessing. Heartbeats were cut with a fixed length of $[-200, 400]$ ms around each R peak. Outliers were detected and removed using DMEAN ($\alpha = 0.5$ and $\beta = 1.5$, with Euclidean distance). For decision, the k-nearest neighbors (kNN) classifier was used with $k = 3$ and cosine distance.

Standard sample wise normalization was performed (following Eq. 1) for all methods except that of Eduardo *et al.* [28], which required $[-1, 1]$ *min-max* normalization (Eq. 2), where x represents the input signal and $\tilde{x}$ the normalized signal.

$$\tilde{x}[g] = \frac{x[g] - \overline{x[g]}}{\sigma(x[g])} \tag{1}$$

$$\tilde{x}[g] = 2\left(\frac{x[g] - \min(x[g])}{\max(x[g]) - \min(x[g])}\right) - 1 \tag{2}$$

3.2 Template update methods

FIFO (first-in-first-out) is the most common strategy and, computationally, is very light. Here, the database is updated using new samples whose score is above or below a threshold (whether the score represents similarity or dissimilarity, respectively), or between two threshold values (discarding previously stored sample) [8, 29]. The score of a new sample can either be output by a classifier, or be a measure of distance or similarity between that sample and the stored templates [30].

In this work, the training data were used to search for threshold values. Among all training samples, 75% were used to train a model, which was used to obtain scores for the remaining data samples. Comparing the scores with several thresholds, the error at each threshold was analyzed (Fig. 2) to find one that simultaneously maximizes true positives and minimizes false positives.

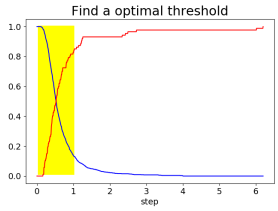


Fig. 2: Illustration of the search for the ideal threshold. The values were chosen near the intersection, inside the yellow zone.

Fixation consists on fixing certain templates, allowing only the remaining stored samples to be updated [31]. In this work, 25, 50, or 75% of the enrollment templates of the individual are fixed, while the rest of the samples are free to be updated. This ensures some initial, labeled information of the subjects remains on the system over time.

An adaptation of this technique was explored. Here, $n + j \times n$ samples were fixed, where $n \in [1, 2, 3]$ is the number of fixed initial templates, and j increases over time. In this work, $j \in [0, 6]$ increased by one at each testing moment ($j \in [0, 6]$), which allowed the system to fix more and more samples over time, thus storing information on the subject's variability over time. In a real system with potentially endless use, the parameters n and j should be carefully chosen to avoid the eventual fixation of the entire template gallery.

4 Experimental Settings

4.1 Dataset

For evaluation, the ECG signals used were from the E-HOL-03-0202-003 database[3] (most commonly designated as E-HOL 24h). This database consists of a study of 202 healthy subjects (only 201 were provided), recorded using three leads at 200 Hz sampling frequency, after an initial resting supine period of 20 minutes. From the available data of 201 subjects, thirteen were discarded due to saturation or unacceptable noise (subjects 1043, 9003, 9005, 9020, 9021, 9022, 9025, 9046, 9061, 9064, 9071, 9082 and 9105), similar to what was done by Labati *et al.* [14]. From each of the remaining 188 subjects, only the lead most closely resembling Lead I ECG was selected, to approximate off-the-person settings.

4.2 Experiments

In order to fit the used data, some changes were introduced to the original methods. On the method from Eduardo *et al.*, the cutoff frequencies of the

[3] THEW. Available on: http://thew-project.org/Database/E-HOL-03-0202-003.html.

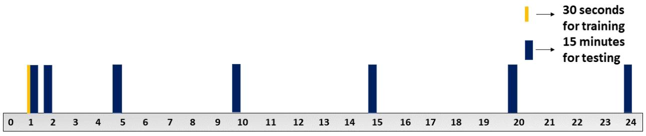


Fig. 3: Schema illustrating the use of each E-HOL record for training and testing (in orange - training segment; in blue - each test segment).

bandpass filter were changed to 1 and 40 Hz, to retrieve important information on higher frequencies; outlier removal was reparametrized with $\alpha = 1.2$ and $\beta = 1.5$. The autoencoder had the topology $[120, 60, 40, 20, 40, 60, 120]$ and was trained using the Adam optimizer with learning rate 0.01. Classification used $k = 1$.

For the method of Belgacem *et al.*, DWT feature extraction was performed in four decomposition levels, due to lower data sampling rate, and $cd4$, $cd3$, $cd2$, and $cd1$ coefficients were used as features.

Data were divided into train and test sets. The training phase used the last 30 seconds (mimicking short enrollments on real-life applications) of the first 60 minutes (avoiding unrealistic calm after the initial resting period) of each subject. Five-second overlap was used to obtain 26 samples from each 30 seconds of training data. To study performance over time, testing was performed over seven time points (see Fig. 3): one immediately after enrollment, another after one hour, and regularly until the end of the records. From each point, from 15 minutes of data, thirty 30 s samples are extracted, and batches are built with one sample from each of the 188 subjects.

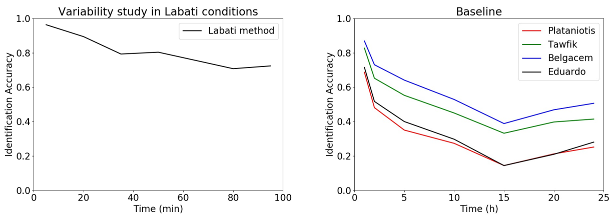
5 Results

5.1 Without Template Update

After implementing, for identification, the method proposed by Labati *et al.* [10] (replicating their evaluation conditions), it was possible to conclude that the ECG signal is not fully permanent over 24 h. However, similarly to what was stated by Labati *et al.*, the results are relatively good over the first two hours (see Fig 4a), although permanence was not verified.

The performance results at each test hour, obtained through the weighted average of the corresponding batches, for the state-of-the-art methods can be found in Fig. 4b. It was found that the performance is mostly acceptable in the first test point, but performance decays significantly over time and variability changes considerably over the day.

Moreover, a minimum around the 15^{th} hour occurs independently of the chosen method. Considering that most of the records start between 8-12 am,



(a) (b)

Fig. 4: Identification performance over time corresponding to (a) the Labati *et al.* method, and (b) the implemented state-of-the-art methods.

after 15 hours the subjects must be sleeping. In this perspective, it appears that the ECG is most different from normal when the subject is asleep.

5.2 With Template Update

Considering the previous results, template update was applied to the methods, in an effort to avoid performance decay over time. Fig. 5 presents the results using the FIFO technique, with diverse thresholds.

For the methods of Plataniotis *et al.* and Eduardo *et al.*, the best results were obtained using two thresholds, respectively, $\{0.3, 0.7\}$ (4.7% accuracy improvement) and $\{0.1, 0.3\}$ (+ 5.7% accuracy), improving all performance results until the 15$^{\text{th}}$ hour. However, for the method of Belgacem *et al.*, the performance worsens with template update after the first two hours (best results obtained when the difference between the highest and second highest scores $\Delta score \in [0.15, 0.3]$). The same was verified for the method of Tawfik *et al.* which, in the first two hours, offered best results with $\Delta score > 0.2$. In general, using two thresholds instead of one offered the best results.

Considering this, it appeared that the Random Forest and MLP classifiers are not suitable for these kinds of template/model update. This was confirmed after a repetition of the evaluation of these methods, with kNN replacing the classifiers (see Fig. 6). With kNN, the template update was able to reduce the performance decay over time, improving accuracy, on average, by 7.9% and 9.2%, respectively, for the methods of Belgacem *et al.* and Tawfik *et al.*

As for the Fixation technique, the obtained results were more promising (see Fig. 7). This template update technique brought performance improvements for all methods. The fixation technique that offered the best results was $j \times 3 + 3$, improving the baseline identification accuracy, on average, by 10.0%.

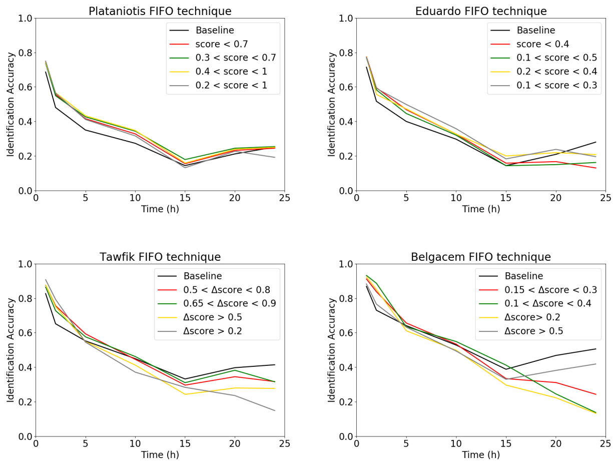


Fig. 5: Comparative FIFO methods with different thresholds using different identification methodologies.

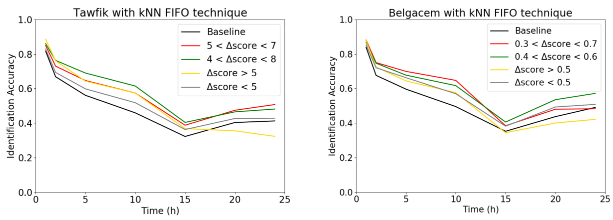


Fig. 6: Results using FIFO update with different thresholds.

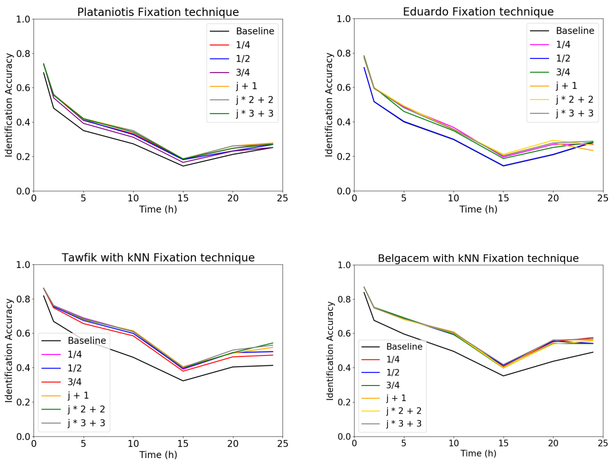


Fig. 7: Results using Fixation update (the corresponding value represents the number of samples that was fixated per subject).

6 Conclusion

This work studied how the ECG variability effects the performance of state-of-the-art biometric algorithms, and how template update could mitigate performance decay over time. The results have shown long-term identification performance in ECG biometrics is generally weak, despite the promising results often presented in the literature.

Template update techniques proved successful in enhancing the long-term performance of state-of-the-art methods, especially when using template fixation techniques. However, further efforts are needed for the study and development of more advanced techniques, with special focus on supervised techniques, so that ECG-based biometric systems can offer reliable performances over long periods.

Acknowledgements

This work was financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalization - COMPETE 2020 Programme and by National Funds through the Portuguese funding agency, FCT – Fundação para a Ciência e a Tecnologia within project "POCI-01-0145-FEDER-030707", and within the PhD grant "SFRH/BD/137720/2018". Data used for this research was provided by the Telemetric and Holter ECG Warehouse of the University of Rochester (THEW), NY.

References

1. S. Prabhakar, S. Pankanti, and A.K. Jain. Biometric recognition: security and privacy concerns. *IEEE Security & Privacy Magazine*, 1(2):33–42, March 2003.
2. A. Hadid, N. Evans, S. Marcel, and Julian F. Biometrics Systems Under Spoofing Attack: An evaluation methodology and lessons learned. *IEEE Signal Processing Magazine*, 32(5):20–30, September 2015.
3. F. Agrafioti, F. M. Bui, and D. Hatzinakos. Secure Telemedicine: Biometrics for Remote and Continuous Patient Verification. *Journal of Computer Networks and Communications*, 2012:1–11, 2012.
4. Z. Akhtar, C. Micheloni, and G. Foresti. Biometric Liveness Detection: Challenges and Research Opportunities. *IEEE Security & Privacy*, 13(5):63–72, September 2015.
5. A.K. Jain, A. Ross, and S. Prabhakar. An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, January 2004.
6. J. R. Pinto, J. S. Cardoso, A. Lourenço, and C. Carreiras. Towards a Continuous Biometric System Based on ECG Signals Acquired on the Steering Wheel. *Sensors*, 17(10):2228, 2017.
7. J. R. Pinto, J. S. Cardoso, and A. Lourenco. Evolution, Current Challenges, and Future Possibilities in ECG Biometrics. *IEEE Access*, 6:34746–34776, 2018.
8. M. Komeili, N. Armanfard, and D. Hatzinakos. Liveness Detection and Automatic Template Updating Using Fusion of ECG and Fingerprint. *IEEE Transactions on Information Forensics and Security*, 13(7):1810–1822, July 2018.
9. B. Freni. *Template Editing and Replacement: novel methods for Biometric Template Selection and Update*. PhD thesis, University of Cagliari, Italy, 2010.
10. R. Labati, R. Sassi, and F. Scotti. ECG biometric recognition: Permanence analysis of QRS signals for 24 hours continuous authentication. In *2013 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 31–36, Guangzhou, China, November 2013. IEEE.
11. A. Rattani. *Adaptive Biometric System based on Template Update Procedures*. PhD thesis, University of Cagliari, 2010.
12. A. Lumini and L. Nanni. A clustering method for automatic biometric template selection. *Pattern Recognition*, 39(3):495–497, March 2006.
13. D. Hutchison and G. Anagnostopoulos. *Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshop, SSPR & SPR 2008, Orlando, USA, December 4-6, 2008. Proceedings*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
14. R. D. Labati, V. Piuri, R. Sassi, F. Scotti, and G. Sforza. Adaptive ECG biometric recognition: a study on re-enrollment methods for QRS signals. In *2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, pages 30–37, Orlando, FL, USA, December 2014. IEEE.
15. S. Y. Chun. Small scale single pulse ECG-based authentication using GLRT that considers T wave shift and adaptive template update with prior information. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pages 3043–3048, Cancun, Mexico, December 2016. IEEE.

16. T. Scheidat, A. Makrushin, and C. Vielhauer. Automatic Template Update Strategies for Biometrics. Technical report, Otto-von-Guericke University of Magdeburg, Germany, 2007.
17. X. Zhu. *Semi-Supervised Learning Literature Survey*. University of Wisconsin-Madison, USA, 2006.
18. A. Blum and S. Chawla. Learning from Labeled and Unlabeled Data using Graph Mincuts. In *Proceedings of International Conference on Machine Learning (ICML-2001)*, 2001.
19. X. Zhu, Z. Ghahramani, and J. Lafferty. Semi-Supervised Learning Using Gaussian Fields and Harmonic Functions. *Twentieth International Conference on Machine Learning (ICML-2003), Washington DC*, page 8, 2003.
20. Z. Zha, T. Mei, J. Wang, Z. Wang, and X. Hua. Graph-based semi-supervised learning with multiple labels. *Journal of Visual Communication and Image Representation*, 20(2):97–103, 2009.
21. M. Belkin, I. Matveeva, and P. Niyogi. Regularization and Semi-supervised Learning on Large Graphs. In David Hutchison, editor, *Learning Theory*, volume 3120, pages 624–638. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.
22. V. Sindhwani, P. Niyogi, M. Belkin, and S. Keerthi. Linear Manifold Regularization for Large Scale Semi-supervised Learning. In *Proceedings of the 22nd ICML Workshop on Learning with Partially Classified Training Data*, page 4, Bonn, Germany, August 2005.
23. O. Chapelle, B. Schölkopf, and A. Zien, editors. *Semi-supervised learning*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass, 2006.
24. Y. Kokkinos and K. Margaritis. Local learning regularization networks for localized regression. *Neural Computing and Applications*, 28(6):1309–1328, June 2017.
25. K. Plataniotis, D. Hatzinakos, and J. Lee. ECG Biometric Recognition Without Fiducial Detection. In *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6, Baltimore, MD, USA, September 2006. IEEE.
26. M. M. Tawfik, H. Selim, and T. Kamal. Human identification using time normalized QT signal and the QRS complex of the ECG. In *2010 7th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP 2010)*, pages 755–759, July 2010.
27. N. Belgacem, A. Nait-Ali, R. Fournier, and F. Bereksi-Reguig. ECG Based Human Authentication using Wavelets and Random Forests. *International Journal on Cryptography and Information Security*, 2(2):1–11, June 2012.
28. A. Eduardo, H. Aidos, and A. Fred. ECG-based Biometrics using a Deep Autoencoder for Feature Learning - An Empirical Study on Transferability:. In *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods*, pages 463–470, Porto, Portugal, 2017.
29. D. P. Coutinho, A. L. N. Fred, and M.A.T. Figueiredo. ECG-Based continuous authentication system using adaptive string matching:. In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*, pages 354–359, Rome, Italy, 2011.
30. A. Lourenço, H. Silva, and A. Fred. Unveiling the Biometric Potential of Finger-Based ECG Signals. *Computational Intelligence and Neuroscience*, 2011:1–8, 2011.
31. J. Guerra-Casanova, C. Sánchez-Ávila, A. S. Sierra, and G. B. del Pozo. Score optimization and template updating in a biometric technique for authentication in mobiles based on gestures. *Journal of Systems and Software*, 84(11):2013–2021, 2011.

References

A. J. Abboud and S. A. Jassim. Biometric templates selection and update using quality measures. *Proceedings of SPIE - The International Society for Optical Engineering*, 8406:840609, May 2012. doi: 10.1117/12.918772. URL http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.918772.

M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas. Biometric authentication based on PCG and ECG signals: present status and future directions. *Signal, Image and Video Processing*, 8(4): 739–751, May 2014. ISSN 1863-1703, 1863-1711. doi: 10.1007/s11760-013-0593-4. URL http://link.springer.com/10.1007/s11760-013-0593-4.

F. Agrafioti, F. M. Bui, and D. Hatzinakos. Secure Telemedicine: Biometrics for Remote and Continuous Patient Verification. *Journal of Computer Networks and Communications*, 2012: 1–11, 2012. ISSN 2090-7141, 2090-715X. doi: 10.1155/2012/924791. URL http://www.hindawi.com/journals/jcnc/2012/924791/.

Z. Akhtar, A. Ahmed, C. E. Erdem, and G. L. Foresti. Biometric template update under facial aging. In *2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, pages 9–15, Orlando, FL, USA, Dec. 2014. IEEE. ISBN 978-1-4799-4533-7. doi: 10.1109/CIBIM.2014.7015437. URL http://ieeexplore.ieee.org/document/7015437/.

Z. Akhtar, C. Micheloni, and G. L. Foresti. Biometric Liveness Detection: Challenges and Research Opportunities. *IEEE Security & Privacy*, 13(5):63–72, Sept. 2015. ISSN 1540-7993. doi: 10.1109/MSP.2015.116. URL http://ieeexplore.ieee.org/document/7310809/.

V. Allen and J. Belina. ECG data compression using the discrete cosine transform (DCT). In *Proceedings Computers in Cardiology*, pages 687–690, Durham, NC, USA, 1992. IEEE Comput. Soc. Press. ISBN 978-0-8186-3552-6. doi: 10.1109/CIC.1992.269340. URL http://ieeexplore.ieee.org/document/269340/.

G. E. Arrobo, C. A. Perumalla, Yang Liu, T. P. Ketterl, R. D. Gitlin, and P. J. Fabri. novel vectorcardiogram system. In *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pages 243–247, Natal, Oct. 2014. IEEE. ISBN 978-1-4799-6644-8. doi: 10.1109/HealthCom.2014.7001848. URL http://ieeexplore.ieee.org/document/7001848/.

D. S. Baim, W. S. Colucci, E. S. Monrad, H. S. Smith, R. F. Wright, A. Lanoue, D. F. Gauthier, B. J. Ransil, W. Grossman, and E. Braunwald. Survival of patients with severe congestive heart failure treated with oral milrinone. *Journal of the American College of Cardiology*, 7(3): 661–670, Mar. 1986. ISSN 07351097. doi: 10.1016/S0735-1097(86)80478-8. URL http://linkinghub.elsevier.com/retrieve/pii/S0735109786804788.

M. Be, A. M Abhishek, T. J, V. K R, and L. M Patnaik. Multimodal Biometric Authentication using ECG and Fingerprint. *International Journal of Computer Applications*, 111(13):33–39, Feb. 2015. ISSN 09758887. doi: 10.5120/19601-1452. URL http://research.ijcaonline.org/volume111/number13/pxc3901452.pdf.

N. Belgacem. ECG Based Human Authentication using Wavelets and Random Forests. *International Journal on Cryptography and Information Security*, 2(2):1–11, June 2012. ISSN 18398626. doi: 10.5121/ijcis.2012.2201. URL http://www.airccse.org/journal/ijcis/papers/2212ijcis01.pdf.

M. Belkin, I. Matveeva, and P. Niyogi. Regularization and Semi-supervised Learning on Large Graphs. In D. Hutchison, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, J. Shawe-Taylor, and Y. Singer, editors, *Learning Theory*, volume 3120, pages 624–638. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. ISBN 978-3-540-22282-8 978-3-540-27819-1. doi: 10.1007/978-3-540-27819-1_43. URL http://link.springer.com/10.1007/978-3-540-27819-1_43.

J. G. Betts, P. Desaix, E. Johnson, J. E. Johnson, O. Korol, D. Kruse, B. Poe, J. Wise, M. D. Womble, K. A. Young, and OpenStax College. *Anatomy & physiology*. OpenStax, 2017. ISBN 978-1-938168-13-0. OCLC: 1005680121.

J. Beutel, editor. *Handbook of medical imaging*. SPIE Press, Bellingham, Wash, 2000. ISBN 978-0-8194-3621-4 978-0-8194-3622-1 978-0-8194-3623-8.

L. Biel, O. Pettersson, L. Philipson, and P. Wide. ECG Analysis: A New Approach in Human Identification. *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, 50 (3):5, 2001.

C. M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. Springer, New York, 2006. ISBN 978-0-387-31073-2.

A. Blum and S. Chawla. Learning from labeled and unlabeled data using graph mincuts. In *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01, pages 19–26, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc. ISBN 1-55860-778-1. URL http://dl.acm.org/citation.cfm?id=645530.757779.

R. Bolle. *Guide to biometrics*. Springer, New York; London, 2011. ISBN 978-1-4419-2305-9. OCLC: 752480927.

Bor-Shyh Lin, W. Chou, Hsing-Yu Wang, Yan-Jun Huang, and Jeng-Shyang Pan. Development of Novel Non-Contact Electrodes for Mobile Electrocardiogram Monitoring System. *IEEE Journal of Translational Engineering in Health and Medicine*, 1, 2013. ISSN 2168-2372. doi: 10.1109/JTEHM.2013.2253598. URL http://ieeexplore.ieee.org/document/6523100/.

J. D. Bronzino and C. R. Company, editors. *The biomedical engineering handbook. Vol. 1: ...* The electrical engineering handbook series. CRC Press [u.a.], Boca Raton, Fla., 2. ed edition, 2000. ISBN 978-0-8493-0461-3 978-3-540-66351-5. OCLC: 76134666.

O. Chapelle, B. Schölkopf, and A. Zien, editors. *Semi-supervised learning*. Adaptive computation and machine learning. MIT Press, Cambridge, Mass, 2006. ISBN 978-0-262-03358-9. OCLC: ocm64898359.

S. Y. Chun. Small scale single pulse ECG-based authentication using GLRT that considers T wave shift and adaptive template update with prior information. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pages 3043–3048, Cancun, Dec. 2016. IEEE. ISBN 978-1-5090-4847-2. doi: 10.1109/ICPR.2016.7900101. URL http://ieeexplore.ieee.org/document/7900101/.

L. Costaridou. Medical Image Analysis Methods. page 552, 2005.

D. P. Coutinho. ECG-BASED CONTINUOUS AUTHENTICATION SYSTEM USING ADAPTIVE STRING MATCHING:. In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*, pages 354–359, Rome, Italy, 2011. SciTePress - Science and and Technology Publications. ISBN 978-989-8425-35-5. doi: 10.5220/0003292003540359. URL http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0003292003540359.

J. P. S. Cunha, B. Cunha, A. S. Pereira, W. Xavier, N. Ferreira, and L. Meireles. Vital-Jacket®: A wearable wireless vital signs monitor for patients' mobility in cardiology and sports. In *Proceedings of the 4th International ICST Conference on Pervasive Computing Technologies for Healthcare*, Munchen, Germany, 2010. IEEE. ISBN 978-963-9799-89-9. doi: 10.4108/ICST.PERVASIVEHEALTH2010.899. URL http://eudl.eu/doi/10.4108/ICST.PERVASIVEHEALTH2010.899.

T. Cutmore. Identifying and reducing noise in psychophysiological recordings. *International Journal of Psychophysiology*, 32(2):129–150, May 1999. ISSN 01678760. doi: 10.1016/S0167-8760(99)00014-8. URL http://linkinghub.elsevier.com/retrieve/pii/S0167876099000148.

H. P. da Silva, A. Lourenco, C. Carreiras, and A. Fred. Off-the-Person Electrocardiography:. In *Proceedings of the International Congress on Cardiovascular Technologies*, pages 99–106, Vilamoura, Algarve, Portugal, 2013. SCITEPRESS - Science and and Technology Publications. ISBN 978-989-8565-78-5. doi: 10.5220/0004647700990106. URL http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0004647700990106.

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. *Computer Methods and Programs in Biomedicine*, 113(2):503–514, Feb. 2014. ISSN 01692607. doi: 10.1016/j.cmpb.2013.11.017. URL https://linkinghub.elsevier.com/retrieve/pii/S0169260713003891.

H. P. da Silva, C. Carreiras, A. Lourenço, A. Fred, R. C. das Neves, and R. Ferreira. Off-the-person electrocardiography: performance assessment and clinical correlation. *Health and Technology*, 4(4):309–318, Apr. 2015. ISSN 2190-7188, 2190-7196. doi: 10.1007/s12553-015-0098-y. URL http://link.springer.com/10.1007/s12553-015-0098-y.

K. Delac and M. Grgic. A SURVEY OF BIOMETRIC RECOGNITION METHODS. In *Proceedings. Elmar-2004. 46th International Symposium on Electronics in Marine*, page 10. IEEE, 2004.

S. Eberz, N. Paoletti, M. Roeschlin, A. Patani, M. Kwiatkowska, and I. Martinovic. Broken Hearted: How To Attack ECG Biometrics. In *Proceedings 2017 Network and Distributed System Security Symposium*, San Diego, CA, 2017a. Internet Society. ISBN 978-1-891562-46-4. doi: 10.14722/ndss.2017.23408. URL https://www.ndss-symposium.org/ndss2017/ndss-2017-programme/broken-hearted-how-attack-ecg-biometrics/.

S. Eberz, K. B. Rasmussen, V. Lenders, and I. Martinovic. Evaluating Behavioral Biometrics for Continuous Authentication: Challenges and Metrics. In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security - ASIA CCS '17*, pages 386–399, Abu Dhabi, United Arab Emirates, 2017b. ACM Press. ISBN 978-1-4503-4944-4. doi: 10.1145/3052973.3053032. URL http://dl.acm.org/citation.cfm?doid=3052973.3053032.

A. Eduardo, H. Aidos, and A. Fred. ECG-based Biometrics using a Deep Autoencoder for Feature Learning - An Empirical Study on Transferability:. In *Proceedings of the 6th International Conference on Pattern Recognition Applications and Methods*, pages 463–470, Porto, Portugal, 2017. SCITEPRESS - Science and Technology Publications. ISBN 978-989-758-222-6. doi: 10.5220/0006195404630470. URL http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0006195404630470.

A. el Ouardighi, A. el Akadi, and D. Aboutajdine. Feature Selection on Supervised Classification Using Wilks Lambda Statistic. In *2007 International Symposium on Computational Intelligence and Intelligent Informatics*, pages 51–55, Agadir, Morocco, Mar. 2007. IEEE. ISBN 978-1-4244-1157-3 978-1-4244-1158-0. doi: 10.1109/ISCIII.2007.367361. URL http://ieeexplore.ieee.org/document/4218394/.

J. Fatisson, V. Oswald, and F. Lalonde. Influence Diagram of Physiological and Environmental Factors Affecting Heart Rate Variability: An Extended Literature Overview. *Heart International*, 11(1):heartint.500023, Jan. 2016. ISSN 1826-1868, 2036-2579. doi: 10.5301/heartint.5000232. URL http://journals.sagepub.com/doi/10.5301/heartint.5000232.

B. Freni. Template Editing and Replacement: novel methods for Biometric Template Selection and Update. phdphd thesis, University of Cagliari, Cagliari (Italy), 2010. <div id="ss3468400"style="width:425px"><objectclassid="clsid:D27CDB6E−AE6D−11cf−96B8−444553540000"codebase="http://download.macromedi 6,0,29,0"width = "425"height = "355" >< paramname = "movie"value = "http : //static.slidesharecdn.com/swf/ssplayer2.swf?doc = freni − phdslides − pdf − 100318113225 − phpapp01&stripped_title = freni − ph − d − slides − pdf"/ >< paramname = "quality"value = "high"/ >< paramname = "menu"value = "false"/ >< paramname = "wmode"value = ""/ >< embedsrc = "http : //static.slidesharecdn.com/swf/ssplayer2.swf?doc = freni − phdslides − pdf − 100318113225 − phpapp01&stripped_title = freni − ph − d − slides − pdf"wmode = ""quality = "high"menu = "false"pluginspage = "http : //www.macromedia.com/go/getflashplayer"type = "application/x − shockwave − flash"width = "425"height = "355" >< /embed >< /object >< /div > .

B. Freni, G. L. Marcialis, and F. Roli. Online and Offline Fingerprint Template Update Using Minutiae: An Experimental Comparison. In F. J. Perales and R. B. Fisher, editors, *Articulated Motion and Deformable Objects*, volume 5098, pages 441–448. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-70516-1 978-3-540-70517-8. doi: 10.1007/978-3-540-70517-8_43. URL http://link.springer.com/10.1007/978-3-540-70517-8_43.

G. Gates. The reduced nearest neighbor rule (Corresp.). *IEEE Transactions on Information Theory*, 18(3):431–433, May 1972. ISSN 0018-9448. doi: 10.1109/TIT.1972.1054809. URL http://ieeexplore.ieee.org/document/1054809/.

R. Giot, C. Rosenberger, and B. Dorizzi. Hybrid template update system for unimodal biometric systems. In *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–7, Arlington, VA, USA, Sept. 2012. IEEE. ISBN 978-1-4673-1385-8 978-1-4673-1384-1 978-1-4673-1383-4. doi: 10.1109/BTAS.2012.6374539. URL http://ieeexplore.ieee.org/document/6374539/.

D. O. Gorodnichy. Evolution and evaluation of biometric systems. In *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pages 1–8, Ottawa, ON, Canada, July 2009. IEEE. ISBN 978-1-4244-3763-4. doi: 10.1109/CISDA.2009.5356531. URL http://ieeexplore.ieee.org/document/5356531/.

P. J. Grother, G. W. Quinn, and P. J. Phillips. Report on the evaluation of 2d still-image face recognition algorithms. Technical Report NIST IR 7709, National Institute of Standards and Technology, Gaithersburg, MD, 2011. URL https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7709.pdf.

M. Guennoun, N. Abbad, J. Talom, S. M. M. Rahman, and K. El-Khatib. Continuous authentication by electrocardiogram data. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*, pages 40–42, Toronto, ON, Canada, Sept. 2009. IEEE. ISBN 978-1-4244-3877-8. doi: 10.1109/TIC-STH.2009.5444466. URL http://ieeexplore.ieee.org/document/5444466/.

J. Guerra-Casanova, C. Sánchez-Ávila, A. de Santos Sierra, and G. B. del Pozo. Score optimization and template updating in a biometric technique for authentication in mobiles based on gestures. *Journal of Systems and Software*, 84(11):2013–2021, Nov. 2011. ISSN 01641212. doi: 10.1016/j.jss.2011.05.059. URL http://linkinghub.elsevier.com/retrieve/pii/S0164121211001427.

I. Guyon, editor. *Feature extraction: foundations and applications*. Number v. 207 in Studies in fuzziness and soft computing. Springer-Verlag, Berlin ; New York, 2006. ISBN 978-3-540-35487-1. doi: 10.1007/978-3-540-35488-8. OCLC: ocm70886217.

A. Hadid, N. Evans, S. Marcel, and J. Fierrez. Biometrics Systems Under Spoofing Attack: An evaluation methodology and lessons learned. *IEEE Signal Processing Magazine*, 32(5):20–30, Sept. 2015. ISSN 1053-5888. doi: 10.1109/MSP.2015.2437652. URL http://ieeexplore.ieee.org/document/7194916/.

T. Hastie, R. Tibshirani, and J. H. Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer series in statistics. Springer, New York, NY, 2nd ed edition, 2009. ISBN 978-0-387-84857-0 978-0-387-84858-7.

M. Heusner, T. Keller, and M. Helmert. Best-Case and Worst-Case Behavior of Greedy Best-First Search. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pages 1463–1470, Stockholm, Sweden, July 2018. International Joint Conferences on Artificial Intelligence Organization. ISBN 978-0-9992411-2-7. doi: 10.24963/ijcai.2018/203. URL https://www.ijcai.org/proceedings/2018/203.

Z. M. Hira and D. F. Gillies. A Review of Feature Selection and Feature Extraction Methods Applied on Microarray Data. *Advances in Bioinformatics*, 2015:1–13, 2015. ISSN 1687-8027, 1687-8035. doi: 10.1155/2015/198363. URL http://www.hindawi.com/journals/abi/2015/198363/.

N. Houmani and S. Garcia-Salicetti. On Hunting Animals of the Biometric Menagerie for Online Signature. *PLOS ONE*, 11(4):e0151691, Apr. 2016. ISSN 1932-6203. doi: 10.1371/journal.pone.0151691. URL http://dx.plos.org/10.1371/journal.pone.0151691.

D. Hutchison, G. C. Anagnostopoulos, M. Georgiopoulos, T. Kanade, T. Kasparis, J. Kittler, J. M. Kleinberg, J. T. Kwok, M. Loog, F. Mattern, J. C. Mitchell, M. Naor, O. Nierstrasz, C. Pandu Rangan, F. Roli, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, N. Vitoria Lobo, and G. Weikum. *Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshop, SSPR & SPR 2008, Orlando, USA, December 4-6, 2008. Proceedings*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-89689-0. URL http://dx.doi.org/10.1007/978-3-540-89689-0. OCLC: 698780606.

ISO. ISO 19092:2008. Technical report, 2008. URL https://www.iso.org/obp/ui/#iso:std:iso:19092:ed-1:v1:en.

S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold. ECG to identify individuals. *Pattern Recognition*, 38(1):133–142, Jan. 2005. ISSN 00313203. doi: 10.1016/j.patcog.2004.05.014. URL http://linkinghub.elsevier.com/retrieve/pii/S0031320304002419.

F. Jager, A. Taddei, G. B. Moody, M. Emdin, G. Antolič, R. Dorn, A. Smrdel, C. Marchesi, and R. G. Mark. Long-term ST database: A reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. *Medical & Biological Engineering & Computing*, 41(2):172–182, Mar. 2003. ISSN 0140-0118, 1741-0444. doi: 10.1007/BF02344885. URL http://link.springer.com/10.1007/BF02344885.

A. Jain, A. Ross, and S. Prabhakar. An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, Jan. 2004. ISSN 1051-8215. doi: 10.1109/TCSVT.2003.818349. URL http://ieeexplore.ieee.org/document/1262027/.

A. K. Jain and A. Kumar. Biometric Recognition: An Overview. In E. Mordini and D. Tzovaras, editors, *Second Generation Biometrics: The Ethical, Legal and Social Context*, volume 11, pages 49–79. Springer Netherlands, Dordrecht, 2012. ISBN 978-94-007-3891-1 978-94-007-3892-8. doi: 10.1007/978-94-007-3892-8_3. URL http://www.springerlink.com/index/10.1007/978-94-007-3892-8_3.

G. James, D. Witten, T. Hastie, and R. Tibshirani, editors. *An introduction to statistical learning: with applications in R*. Number 103 in Springer texts in statistics. Springer, New York, 2013. ISBN 978-1-4614-7137-0. OCLC: ocn828488009.

J. L. Jameson, editor. *Harrisons principles of internal medicine*. McGraw-Hill Education, New York, twentieth edition edition, 2018. ISBN 978-1-259-64404-7 978-1-259-64400-9.

L. Johannesen, J. Vicente, J. Mason, C. Erato, C. Sanabria, K. Waite-Labott, M. Hong, J. Lin, P. Guo, A. Mutlib, J. Wang, W. Crumb, K. Blinova, D. Chan, J. Stohlman, J. Florian, M. Ugander, N. Stockbridge, and D. Strauss. Late sodium current block for drug-induced long QT syndrome: Results from a prospective clinical trial. *Clinical Pharmacology & Therapeutics*, 99(2):214–223, Feb. 2016. ISSN 00099236. doi: 10.1002/cpt.205. URL http://doi.wiley.com/10.1002/cpt.205.

W.-H. Jung and S.-G. Lee. ECG Identification Based on Non-Fiducial Feature Extraction Using Window Removal Method. *Applied Sciences*, 7(11):1205, Nov. 2017. ISSN 2076-3417. doi: 10.3390/app7111205. URL http://www.mdpi.com/2076-3417/7/11/1205.

L. P. Kaelbling, M. L. Littman, and A. W. Moore. Reinforcement Learning: A Survey. *Journal of Artificial Intelligence Research*, 4:237–285, May 1996. ISSN 1076-9757. doi: 10.1613/jair.301. URL https://jair.org/index.php/jair/article/view/10166.

M. Karafiat and L. Burget. Using Smoothed Heteroscedastic Linear Discriminant Analysis in Large Vocabulary Continuous Speech Recognition System. page 8, 2005.

N. M. Kesto. Electrocardiography Circuit Design. page 10, 2013.

Y. Kokkinos and K. G. Margaritis. Local learning regularization networks for localized regression. *Neural Computing and Applications*, 28(6):1309–1328, June 2017. ISSN 0941-0643, 1433-3058. doi: 10.1007/s00521-016-2569-0. URL http://link.springer.com/10.1007/s00521-016-2569-0.

M. Komeili, N. Armanfard, and D. Hatzinakos. Liveness Detection and Automatic Template Updating Using Fusion of ECG and Fingerprint. *IEEE Transactions on Information Forensics and Security*, 13(7):1810–1822, July 2018. ISSN 1556-6013, 1556-6021. doi: 10.1109/TIFS.2018.2804890. URL http://ieeexplore.ieee.org/document/8288857/.

R. D. Labati, R. Sassi, and F. Scotti. ECG biometric recognition: Permanence analysis of QRS signals for 24 hours continuous authentication. In *2013 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 31–36, Guangzhou, China, nov 2013. IEEE. ISBN 978-1-4673-5593-3. doi: 10.1109/WIFS.2013.6707790. URL http://ieeexplore.ieee.org/document/6707790/.

R. D. Labati, V. Piuri, R. Sassi, F. Scotti, and G. Sforza. Adaptive ECG biometric recognition: a study on re-enrollment methods for QRS signals. In *2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, pages 30–37, Orlando, FL, USA, Dec. 2014. IEEE. ISBN 978-1-4799-4533-7. doi: 10.1109/CIBIM.2014.7015440. URL http://ieeexplore.ieee.org/document/7015440/.

P. Laguna, R. Mark, A. Goldberg, and G. Moody. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. In *Computers in Cardiology 1997*, pages 673–676, Lund, Sweden, 1997. IEEE. ISBN 978-0-7803-4445-7. doi: 10.1109/CIC.1997.648140. URL http://ieeexplore.ieee.org/document/648140/.

Y. Lee, J. J. Filliben, R. J. Micheals, and P. Jonathon Phillips. Sensitivity analysis for biometric systems: A methodology based on orthogonal experiment designs. *Computer Vision and Image Understanding*, 117(5):532–550, May 2013. ISSN 10773142. doi: 10.1016/j.cviu.2013.01.003. URL https://linkinghub.elsevier.com/retrieve/pii/S1077314213000040.

C. Levkov, G. Mihov, R. Ivanov, I. Daskalov, I. Christov, and I. Dotsinsky. Removal of power-line interference from the ECG: a review of the subtraction procedure. *BioMedical Engineering OnLine*, 2005.

Y. Li, J. Yin, E. Zhu, C. Hu, and H. Chen. Score Based Biometric Template Selection and Update. In *2008 Second International Conference on Future Generation Communication and Networking*, pages 35–40, Hainan, China, Dec. 2008. IEEE. ISBN 978-0-7695-3431-2. doi: 10.1109/FGCN.2008.208. URL http://ieeexplore.ieee.org/document/4734275/.

Y. Lin and M. Sriyudthsak. Design and Development of Standard 12-Lead ECG Data Acquisition and Monitoring System. *Procedia Computer Science*, 86:136–139, 2016. ISSN 18770509. doi: 10.

1016/j.procs.2016.05.034. URL https://linkinghub.elsevier.com/retrieve/pii/S1877050916303702.

W. Louis, M. Komeili, and D. Hatzinakos. Continuous Authentication Using One-Dimensional Multi-Resolution Local Binary Patterns (1dmrlbp) in ECG Biometrics. *IEEE Transactions on Information Forensics and Security*, 11(12):2818–2832, Dec. 2016. ISSN 1556-6013, 1556-6021. doi: 10.1109/TIFS.2016.2599270. URL http://ieeexplore.ieee.org/document/7539682/.

A. Lourenço, H. Silva, and A. Fred. Unveiling the Biometric Potential of Finger-Based ECG Signals. *Computational Intelligence and Neuroscience*, 2011:1–8, 2011. ISSN 1687-5265, 1687-5273. doi: 10.1155/2011/720971. URL http://www.hindawi.com/journals/cin/2011/720971/.

A. Lumini and L. Nanni. A clustering method for automatic biometric template selection. *Pattern Recognition*, 39(3):495–497, Mar. 2006. ISSN 00313203. doi: 10.1016/j.patcog.2005.11.004. URL http://linkinghub.elsevier.com/retrieve/pii/S0031320305004188.

K. Malasri and L. Wang. Design and Implementation of a SecureWireless Mote-Based Medical Sensor Network. *Sensors*, 9(8):6273–6297, Aug. 2009. ISSN 1424-8220. doi: 10.3390/s90806273. URL http://www.mdpi.com/1424-8220/9/8/6273.

C. Manapragada, G. I. Webb, and M. Salehi. Extremely fast decision tree. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, pages 1953–1962, New York, NY, USA, 2018. ACM. ISBN 978-1-4503-5552-0. doi: 10.1145/3219819.3220005. URL http://doi.acm.org/10.1145/3219819.3220005.

G. L. Marcialis, A. Rattani, and F. Roli. Biometric Template Update: An Experimental Investigation on the Relationship between Update Errors and Performance Degradation in Face Verification. In N. da Vitoria Lobo, T. Kasparis, F. Roli, J. T. Kwok, M. Georgiopoulos, G. C. Anagnostopoulos, and M. Loog, editors, *Structural, Syntactic, and Statistical Pattern Recognition*, volume 5342, pages 684–693. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-89688-3 978-3-540-89689-0. doi: 10.1007/978-3-540-89689-0_72. URL http://link.springer.com/10.1007/978-3-540-89689-0_72.

G. L. Marcialis, L. Didaci, A. Pisano, E. Granger, and F. Roli. Why template self-update should work in biometric authentication systems? In *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, pages 1086–1091, Montreal, QC, Canada, July 2012. IEEE. ISBN 978-1-4673-0382-8 978-1-4673-0381-1 978-1-4673-0380-4. doi: 10.1109/ISSPA.2012.6310451. URL http://ieeexplore.ieee.org/document/6310451/.

P. Melillo, R. Izzo, A. Orrico, P. Scala, M. Attanasio, M. Mirra, N. De Luca, and L. Pecchia. Automatic Prediction of Cardiovascular and Cerebrovascular Events Using Heart Rate Variability Analysis. *PLOS ONE*, 10(3):e0118504, Mar. 2015. ISSN 1932-6203. doi: 10.1371/journal.pone.0118504. URL https://dx.plos.org/10.1371/journal.pone.0118504.

M. Merone, P. Soda, M. Sansone, and C. Sansone. ECG databases for biometric systems: A systematic review. *Expert Systems with Applications*, 67:189–202, Jan. 2017. ISSN 09574174. doi: 10.1016/j.eswa.2016.09.030. URL https://linkinghub.elsevier.com/retrieve/pii/S0957417416305140.

G. Moody, A. Goldberger, S. McClennen, and S. Swiryn. Predicting the onset of paroxysmal atrial fibrillation: the Computers in Cardiology Challenge 2001. In *Computers in Cardiology 2001. Vol.28 (Cat. No.01CH37287)*, pages 113–116, Rotterdam, Netherlands, 2001. IEEE. ISBN 978-0-7803-7266-5. doi: 10.1109/CIC.2001.977604. URL http://ieeexplore.ieee.org/document/977604/.

K. Niinuma and A. K. Jain. Continuous user authentication using temporal information. *Proceedings of SPIE - The International Society for Optical Engineering*, 04 2010. doi: 10.1117/12.847886.

J. Paone and P. J. Flynn. On the consistency of the biometric menagerie for irises and iris matchers. In *2011 IEEE International Workshop on Information Forensics and Security*, pages 1–6, Iguacu Falls, Brazil, Nov. 2011. IEEE. ISBN 978-1-4577-1019-3 978-1-4577-1017-9 978-1-4577-1018-6. doi: 10.1109/WIFS.2011.6123158. URL http://ieeexplore.ieee.org/document/6123158/.

T. Penzel, G. Moody, R. Mark, A. Goldberger, and J. Peter. The apnea-ECG database. In *Computers in Cardiology 2000. Vol.27 (Cat. 00CH37163)*, pages 255–258, Cambridge, MA, USA, 2000. IEEE. ISBN 978-0-7803-6557-5. doi: 10.1109/CIC.2000.898505. URL http://ieeexplore.ieee.org/document/898505/.

S. Petrutiu, A. V. Sahakian, and S. Swiryn. Abrupt changes in fibrillatory wave characteristics at the termination of paroxysmal atrial fibrillation in humans. *EP Europace*, 9(7):466–470, July 2007. ISSN 1532-2092, 1099-5129. doi: 10.1093/europace/eum096. URL http://academic.oup.com/europace/article/9/7/466/553888/Abrupt-changes-in-fibrillatory-wave.

J. Pinto, J. Cardoso, A. Lourenço, and C. Carreiras. Towards a Continuous Biometric System Based on ECG Signals Acquired on the Steering Wheel. *Sensors*, 17(10):2228, Sept. 2017. ISSN 1424-8220. doi: 10.3390/s17102228. URL http://www.mdpi.com/1424-8220/17/10/2228.

K. N. Plataniotis, D. Hatzinakos, and J. K. M. Lee. ECG Biometric Recognition Without Fiducial Detection. In *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6, Baltimore, MD, USA, Sept. 2006. IEEE. ISBN 978-1-4244-0486-5 978-1-4244-0487-2. doi: 10.1109/BCC.2006.4341628. URL http://ieeexplore.ieee.org/document/4341628/.

N. Poh, S. Bengio, and A. Ross. Revisiting Doddington's Zoo: A Systematic Method to Assess User-dependent Variabilities. In *Multimodal User Authentication(MMUA)*, page 7, 2006.

S. Prabhakar, S. Pankanti, and A. Jain. Biometric recognition: security and privacy concerns. *IEEE Security & Privacy Magazine*, 1(2):33–42, Mar. 2003. ISSN 1540-7993. doi: 10.1109/MSECP.2003.1193209. URL http://ieeexplore.ieee.org/document/1193209/.

C. Pummer. Continuous Biometric Authentication using Electrocardiographic (ECG) Data. Master's thesis, Hagenberg, 2016.

G. R. Doddington, W. Liggett, A. Martin, M. Przybocki, and D. Reynolds. Sheep, goats, lambs and wolves: a statistical analysis of speaker performance in the nist 1998 speaker recognition evaluation. 01 1998.

S. Raschka and V. Mirjalili. *Python machine learning: machine learning and deep learning with Python, scikit-learn, and TensorFlow*. Expert insight. Packt Publishing Ltd, Birmingham Mumbai, second edition edition, 2017. ISBN 978-1-78712-593-3. OCLC: 1023625815.

A. Rattani, G. L. Marcialis, and F. Roli. Biometric template update using the graph mincut algorithm : A case study in face verification. In *2008 Biometrics Symposium*, pages 23–28, Tampa, FL, Sept. 2008. IEEE. ISBN 978-1-4244-2566-2. doi: 10.1109/BSYM.2008.4655518. URL http://ieeexplore.ieee.org/document/4655518/.

A. Rattani, F. Roli, M. Tistarelli, and D. G. L. Marcialis. Adaptive Biometric System based on Template Update Procedures. Master's thesis, University of Cagliari, Italy, Cagliari, 2010.

J. Ribeiro Pinto, J. S. Cardoso, and A. Lourenço. Evolution, current challenges, and future possibilities in ecg biometrics. *IEEE Access*, 6:34746–34776, 2018. ISSN 2169-3536. doi: 10.1109/ACCESS.2018.2849870.

G. Ritter, H. Woodruff, S. Lowry, and T. Isenhour. An algorithm for a selective nearest neighbor decision rule (Corresp.). *IEEE Transactions on Information Theory*, 21(6):665–669, Nov. 1975. ISSN 0018-9448. doi: 10.1109/TIT.1975.1055464. URL http://ieeexplore.ieee.org/document/1055464/.

F. Roli, L. Didaci, and G. L. Marcialis. Template Co-update in Multimodal Biometric Systems. In *Advances in Biometrics*, volume 4642, pages 1194–1202. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-74548-8 978-3-540-74549-5. doi: 10.1007/978-3-540-74549-5_124. URL http://link.springer.com/10.1007/978-3-540-74549-5_124.

A. C. Rossi and C. Lueje. ISO/IEC JTC 1/SC 37. Technical report, 2002. URL https://www.iso.org/committee/313770.html.

T. Scheidat, A. Makrushin, and C. Vielhauer. Automatic Template Update Strategies for Biometrics. Technical report, Otto-von-Guericke University of Magdeburg, Germany, 2007.

R. J. A. Schijvenaars. *Intra-individual variability of the electrocardiogram: assessment and exploitation in computerized ECG analysis*. PhD thesis, s.n.], S.l., 2000. OCLC: 67496040.

R. R. Seeley, P. Tate, and T. D. Stephens. *Anatomia & Fisiologia*. Mosby, 1991. ISBN 978-972-8930-62-2.

T. Sim, S. Zhang, R. Janakiraman, and S. Kumar. Continuous Verification Using Multimodal Biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):687–700, Apr. 2007. ISSN 0162-8828, 2160-9292. doi: 10.1109/TPAMI.2007.1010. URL http://ieeexplore.ieee.org/document/4107572/.

V. Sindhwani and P. Niyogi. Linear manifold regularization for large scale semi-supervised learning. In *Proc. of the 22nd ICML Workshop on Learning with Partially Classified Training Data*, 2005.

B. Singh, P. Singh, and S. Budhiraja. Various Approaches to Minimise Noises in ECG Signal: A Survey. In *2015 Fifth International Conference on Advanced Computing & Communication Technologies*, pages 131–137, Haryana, India, Feb. 2015. IEEE. ISBN 978-1-4799-8488-6. doi: 10.1109/ACCT.2015.87. URL http://ieeexplore.ieee.org/document/7079066/.

F. Song, D. Mei, and H. Li. Feature Selection Based on Linear Discriminant Analysis. In *2010 International Conference on Intelligent System Design and Engineering Application*, pages 746–749, Changsha, Hunan, China, Oct. 2010. IEEE. ISBN 978-1-4244-8333-4. doi: 10.1109/ISDEA.2010.311. URL http://ieeexplore.ieee.org/document/5743287/.

F. Sufi and I. Khalil. Faster person identification using compressed ECG in time critical wireless telecardiology applications. *Journal of Network and Computer Applications*, 34(1):282–293, Jan. 2011. ISSN 10848045. doi: 10.1016/j.jnca.2010.07.004. URL http://linkinghub.elsevier.com/retrieve/pii/S1084804510001268.

M. Sugiyama. Dimensionality reduction of multimodal labeled data by local fisher discriminant analysis. *J. Mach. Learn. Res.*, 8:1027–1061, May 2007. ISSN 1532-4435. URL http://dl.acm.org/citation.cfm?id=1248659.1248694.

B. Taji, S. Shirmohammadi, V. Groza, and I. Batkin. Impact of Skin–Electrode Interface on Electrocardiogram Measurements Using Conductive Textile Electrodes. *IEEE Transactions on Instrumentation and Measurement*, 63(6):1412–1422, June 2014. ISSN 0018-9456, 1557-9662. doi: 10.1109/TIM.2013.2289072. URL http://ieeexplore.ieee.org/document/6670113/.

R. Tan and M. Perkowski. Toward Improving Electrocardiogram (ECG) Biometric Verification using Mobile Sensors: A Two-Stage Classifier Approach. *Sensors*, 17(2):410, Feb. 2017. ISSN 1424-8220. doi: 10.3390/s17020410. URL http://www.mdpi.com/1424-8220/17/2/410.

M. M. Tawfik, H. Selim, and T. Kamal. Human identification using time normalized qt signal and the qrs complex of the ecg. In *2010 7th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP 2010)*, pages 755–759, July 2010.

M. N. Teli, J. R. Beveridge, P. J. Phillips, G. H. Givens, D. S. Bolme, and B. A. Draper. Biometric zoos: Theory and experimental evidence. In *2011 International Joint Conference on Biometrics (IJCB)*, pages 1–8, Washington, DC, USA, Oct. 2011. IEEE. ISBN 978-1-4577-1359-0 978-1-4577-1358-3 978-1-4577-1357-6. doi: 10.1109/IJCB.2011.6117479. URL http://ieeexplore.ieee.org/document/6117479/.

G. J. Tortora and B. Derrickson. *Princípios de Anatomia e Fisiologia (14a. ed.)*. Grupo Gen - Guanabara Koogan, Rio de Janeiro, 2016. ISBN 978-85-277-2885-0. URL http://public.eblib.com/choice/publicfullrecord.aspx?p=4883432. OCLC: 992565352.

U. Uludag, A. Ross, and A. Jain. Biometric template selection and update: a case study in fingerprints. *Pattern Recognition*, 37(7):1533–1542, July 2004. ISSN 00313203. doi: 10.1016/j.patcog.2003.11.012. URL http://linkinghub.elsevier.com/retrieve/pii/S0031320304000081.

C. van Ravenswaaij-Arts, L. Kollee, J. Hopman, G. Stoelinga, and H. P van Geijn. Heart rate variability. *Annals of internal medicine*, 118:436–47, 04 1993.

J. Vicente, L. Johannesen, J. W. Mason, W. J. Crumb, E. Pueyo, N. Stockbridge, and D. G. Strauss. Comprehensive T wave Morphology Assessment in a Randomized Clinical Study of Dofetilide, Quinidine, Ranolazine, and Verapamil. *Journal of the American Heart Association*, 4(4), Apr. 2015. ISSN 2047-9980, 2047-9980. doi: 10.1161/JAHA.114.001615. URL https://www.ahajournals.org/doi/10.1161/JAHA.114.001615.

S. Wahabi, S. Pouryayevali, S. Hari, and D. Hatzinakos. On Evaluating ECG Biometric Systems: Session-Dependence and Body Posture. *IEEE Transactions on Information Forensics and Security*, 9(11):2002–2013, Nov. 2014. ISSN 1556-6013, 1556-6021. doi: 10.1109/TIFS.2014.2360430. URL http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6910235.

H. Wang, H. Fang, L. Xing, and M. Chen. An Integrated Biometric-Based Security Framework Using Wavelet-Domain HMM in Wireless Body Area Networks (WBAN). In *2011 IEEE International Conference on Communications (ICC)*, pages 1–5, Kyoto, Japan, June 2011. IEEE. ISBN 978-1-61284-232-5. doi: 10.1109/icc.2011.5962757. URL http://ieeexplore.ieee.org/document/5962757/.

J. Wayman, editor. *Biometric systems: technology, design, and performance evaluation*. Springer, London, 2005. ISBN 978-1-85233-596-0.

D. Widjaja, C. Varon, A. Dorado, J. A. K. Suykens, and S. Van Huffel. Application of Kernel Principal Component Analysis for Single-Lead-ECG-Derived Respiration. *IEEE Transactions on Biomedical Engineering*, 59(4):1169–1176, Apr. 2012. ISSN 0018-9294, 1558-2531. doi: 10.1109/TBME.2012.2186448. URL http://ieeexplore.ieee.org/document/6144719/.

L. Wieclaw, Y. Khoma, P. Fałat, D. Sabodashko, and V. Herasymenko. Biometrie identification from raw ecg signal using deep learning techniques. In *2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, volume 1, pages 129–133, Sep. 2017. doi: 10.1109/IDAACS.2017.8095063.

D. L. Wilson. Asymptotic Properties of Nearest Neighbor Rules Using Edited Data. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-2(3):408–421, July 1972. ISSN 0018-9472, 2168-2909. doi: 10.1109/TSMC.1972.4309137. URL http://ieeexplore.ieee.org/document/4309137/.

Xudong Jiang and W. Ser. Online fingerprint template improvement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8):1121–1126, Aug. 2002. ISSN 0162-8828. doi: 10.1109/TPAMI.2002.1023807. URL http://ieeexplore.ieee.org/document/1023807/.

N. Yadav, M. Duhan, and A. Rose. Biometric Human Recognition using ECG Signals. *IARJSET*, 4(6):168–171, June 2017. ISSN 23938021. doi: 10.17148/IARJSET.2017.4630. URL http://www.iarjset.com/upload/2017/june-17/IARJSET%2030.pdf.

N. Yager and T. Dunstone. The Biometric Menagerie. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(2):220–230, Feb. 2010. ISSN 0162-8828. doi: 10.1109/TPAMI.2008.291. URL http://ieeexplore.ieee.org/document/4711054/.

S. N. Yanushkevich and A. V. Shmerko. Fundamentals of biometric system design: New course for electrical, computer, and software engineering students. In *2009 Symposium on Bio-inspired Learning and Intelligent Systems for Security*, pages 3–8, Aug 2009. doi: 10.1109/BLISS.2009.27.

C. Ye, M. T. Coimbra, and B. V. K. V. Kumar. Investigation of human identification using two-lead Electrocardiogram (ECG) signals. In *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–8, Washington, DC, USA, Sept. 2010. IEEE. ISBN 978-1-4244-7581-0. doi: 10.1109/BTAS.2010.5634478. URL http://ieeexplore.ieee.org/document/5634478/.

T. W. Yeoh, F. Daolio, H. E. Aguirre, and K. Tanaka. On the effectiveness of feature selection methods for gait classification under different covariate factors. *Applied Soft Computing*, 61:42–57, Dec. 2017. ISSN 15684946. doi: 10.1016/j.asoc.2017.07.041. URL https://linkinghub.elsevier.com/retrieve/pii/S1568494617304611.

Z.-J. Zha, T. Mei, J. Wang, Z. Wang, and X.-S. Hua. Graph-based semi-supervised learning with multiple labels. *Journal of Visual Communication and Image Representation*, 20(2):97–103, Feb. 2009. ISSN 10473203. doi: 10.1016/j.jvcir.2008.11.009. URL `http://linkinghub.elsevier.com/retrieve/pii/S1047320308001144`.

D. Zhang and A. K. Jain, editors. *Biometric authentication: first international conference, ICBA 2004, Hong Kong, China, July 15-17, 2004: proceedings*. Number 3072 in Lecture notes in computer science. Springer, Berlin ; New York, 2004. ISBN 978-3-540-22146-3.

X. Zhu. Semi-supervised learning literature survey. *Comput Sci, University of Wisconsin-Madison*, 2, 07 2008.

X. Zhu, Z. Ghahramani, and J. D. Lafferty. Semi-supervised learning using gaussian fields and harmonic functions. volume 3, pages 912–919, 01 2003.